

Class-E Power Amplifier

By

Richard Kubowicz

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science.
Graduate Department of Electrical and Computer Engineering
University of Toronto



© Copyright by Richard Kubowicz, 2000



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-53437-5

Canada

For Mom, Dad, and Vicky.
For helping me get through it all.

Class-E Power Amplifier

Master of Applied Science, 2000

Richard Kubowicz

Department of Electrical and Computer Engineering

University of Toronto

ABSTRACT

This thesis presents the design and implementation of a Class-E power amplifier implemented in 0.35 μm CMOS. The rationale behind the selection of the final topology is discussed in detail. Comparisons are also drawn to other potential architectures used in practice.

Design details of both gain and output stages are also presented. Theoretical work has been refined for use in a design setting. Many practical high-power RF design issues have also been addressed and discussed in detail. A customized test fixture was also devised to extract optimal performance from the amplifier.

The final amplifier implemented was a fully differential device operating at 1.88 GHz. In general, the amplifier achieved its intended design specifications producing 185 mW with a power added efficiency of 38%.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to Professor C.A.T. Salama. His guidance, support and encouragement made this work possible.

I would also like to thank the many people of the VRG group: Jaro Pristupa, for his mastery of a UNIX terminal, Dana Reem, Richard Barber, Anthoula Vlahakis, and Milena Khazak all made the ride a little smoother.

School is built on the strong foundation of good people. The students and professors at the University of Toronto's Electrical Engineering department are what makes this school great. Although there are literally too many people to mention, a number deserve special recognition. Namely these people are, Merhdad Ramezani, Sameh Khalil, John Ren, Sotoudeh Hamedi-Hagh, An Wei, Jin Heng, and Song Ye. Last but not least, my "partners-in-crime," James Maligeorgos, Vasilis Papanikolaou, and Tyler Paradis made me realize why I get out of bed everyday.

Of course, what would an acknowledgement be without thanking my love Vicky Mattucci. Vicky was always there for me when I needed her. She is indeed a special lady.

Finally, the financial support from Mirconet, Gennum, Mitel, Nortel Networks, and PMC Sierra made this work possible.

TABLE OF CONTENTS

CHAPTER 1

Introduction	1
1.1 Conventional Modern Transceiver Configuration	2
1.2 Frequency Standards	4
1.3 Modulation Standards	5
1.4 Transmit Power Standards	6
1.5 Power Amplifier Options	7
1.5.1 Class B	7
1.5.2 Class C	8
1.5.3 Class D	9
1.5.4 Class E	10
1.5.5 Class F	10
1.5.6 Optimum Choice for a Power Amplifier Architecture	12
1.6 Previous Work	12
1.7 Objectives and Outline of the Thesis	13
1.8 References	15

CHAPTER 2

Class-E Power Amplifier Design	16
2.1 Introduction	16
2.2 Class E Theory of Operation	18
2.3 Class-E Amplifier Design	23
2.4 Differential Architecture	28
2.5 Bondwire Inductors	31
2.6 Source Impedances	32
2.7 Gain Stage Design	35
2.8 Load Pull Analysis	38
2.9 Impedance Matching	40

2.9.1	<i>Spiral Inductor Design</i>	40
2.9.2	<i>Input Matching</i>	45
2.9.3	<i>Output Matching</i>	46
2.10	Thermal Modeling	46
2.11	Final Amplifier Schematic	49
2.12	Theoretical Performance	50
2.13	Summary	53
2.14	References	54

CHAPTER 3

Class-E Power Amplifier Layout		57
3.1	Introduction	57
3.2	Transistor Layout	57
3.3	Capacitor Layout	60
3.4	Grounding Vias	60
3.5	Inductor Layout	62
3.6	Bondpad Layout	63
3.7	RF Signal Paths	64
3.8	Ground Plane Considerations	64
3.9	Final Implementation	65
3.10	Conclusions	65
3.11	References	67

CHAPTER 4

Experimental results		68
4.1	Experimental Implementation	68
4.2	DC Testing	69
4.3	Low Power RF Testing	69
4.4	Low Power Test Results	70
4.5	High Power RF Testing	72
4.6	High Power Test Results	76
4.7	Comparison With Previous Work	82

4.8	Conclusions	82
4.9	References	83

CHAPTER 5

Conclusions	84
--------------------	-----------

LIST OF FIGURES

Figure 1.1:	Typical Present-day Transceiver Configuration _____	2
Figure 1.2:	Generic Configuration for Class B, C, and E Amplifiers _____	7
Figure 1.3:	Class-B Waveforms _____	8
Figure 1.4:	Class-C Waveforms _____	8
Figure 1.5:	Class-D Amplifier _____	9
Figure 1.6:	Class-E Amplifier Waveforms _____	10
Figure 1.7:	Class-F Amplifier Design _____	11
Figure 1.8:	Waveforms for Class-F Amplification _____	11
Figure 2.1:	Class-E Resonator [1.7] _____	17
Figure 2.2:	Simplified Drain Resonant Circuit _____	18
Figure 2.3:	Drain Voltage Waveforms _____	19
Figure 2.4:	Simplified Drain Resonant Circuit _____	20
Figure 2.5:	Drain Capacitance and On-Resistance Proportionality _____	25
Figure 2.6:	Drain Resonant Waveform _____	26
Figure 2.7:	Relationship of Drain Capacitance to Channel Width _____	27
Figure 2.8:	Final Amplifier Waveforms _____	28
Figure 2.9:	Substrate Current Effects on Receiver Circuit [13] _____	29
Figure 2.10:	Comparison of Single-Ended and Differential Substrate Currents _____	29
Figure 2.11:	Mutual Inductance Cancellation _____	30
Figure 2.12:	Test Fixture Ground Plane Configuration _____	33
Figure 2.13:	Ground Plane For Half Circuit _____	34
Figure 2.14:	Effects of Non-Ideal Grounding _____	35
Figure 2.15:	Typical Load Pull Power Contours _____	40
Figure 2.16:	Illustration of Spiral Inductor Modeling Parameters _____	41
Figure 2.17:	Single-Ended Pi Model for CMOS Spiral Inductor _____	41
Figure 2.18:	Spiral Inductor Layout as Simulated in EEsof _____	43

Figure 2.19: Comparison of Empirical Inductor Model and Electromagnetic Analysis	44
Figure 2.20: Comparison of Compensated Empirical and Electromagnetic Inductor Analysis	44
Figure 2.21: S-parameter Sweep of the Amplifier Input Reflection	45
Figure 2.22: Input Reflection and Conjugate Input Match	46
Figure 2.23: Simple Thermal Packaging Model	47
Figure 2.24: Final Amplifier Schematic	50
Figure 2.25: Simulations for Varying Loads Reflections for the Differential Amplifier	51
Figure 2.26: Predicted Performance Versus Frequency	52
Figure 2.27: Predicted Performance Versus Supply Voltage	53
Figure 3.1: Transistor Unit Cell Layout	58
Figure 3.2: Unit Cell Array for Output Device	59
Figure 3.3: Interstage Coupling Capacitor	60
Figure 3.4: Grounding Via Placement	61
Figure 3.5: Input Matching Spiral Inductor	63
Figure 3.6: Comparison of DC and RF Bondpads	64
Figure 3.7: Final Amplifier Layout	65
Figure 4.1: 1.9 GHz Amplifier Photomicrograph	68
Figure 4.2: 1.6 GHz Low Power Spectrum	71
Figure 4.3: 1.8 GHz Low Power Spectrum	71
Figure 4.4: 2 GHz Low Power Spectrum	72
Figure 4.5: Schematic Representation of Chip-On-Board Test Fixture	74
Figure 4.6: Implemented Design Showing Heatsink and Launches	75
Figure 4.7: Chip-On-Board Calibration Structure	76
Figure 4.8: Input Reflection Coefficient	77
Figure 4.9: Input Reflection Magnitude	77
Figure 4.10: Load Pull Analysis Test Setup	78
Figure 4.11: Output Power vs. Supply Voltage	80
Figure 4.12: Efficiency vs. Supply Voltage	80

Figure 4.13: Output Power vs. Frequency _____ 81
Figure 4.14: Efficiency vs. Frequency _____ 81

LIST OF TABLES

Table 1.1:	Modulation Standards Used for PCS Band Communications	5
Table 1.2:	Previous Work on Class-E Power Amplifiers for Wireless Transceivers	13
Table 2.1:	0.35 μm CMOS Process Inductor Design Parameters	43
Table 4.1:	Comparison to Previous Work	82

CHAPTER 1

Introduction

Presently, the wireless telecommunications industry is on the verge of an era of explosive growth. Recent advancements in key technologies have combined to make microwave communications devices operating at low gigahertz frequencies cheap enough so that they can be afforded by the general public. In the very recent past, only people using cellular phones for business applications could afford to buy them. Today, it is possible for the average consumer to purchase a high performance microwave communicator for the price of a week's worth of groceries. As recently as a decade ago, this feat wouldn't have even been conceivable [1].

Advancements in battery-storage and semiconductor technologies have combined to foster an unprecedented level of growth in the wireless industry. The lion's share of this growth can be primarily attributed to the semiconductors used to build a wireless appliance. These new devices have cost and performance figures that are orders of magnitude better than their predecessors.

The ideal goal in the cellular phone industry is to integrate all of the various components used in a cellular phone onto one chip. The power consumption of this ideal chip would also be low enough to allow it to be driven by a single small battery [2]. Given the current distribution of cellular towers within major metropolitan areas, transmit powers of 100 mW are often necessary to achieve good communication links. In these instances, a battery such as this would become quickly depleted.

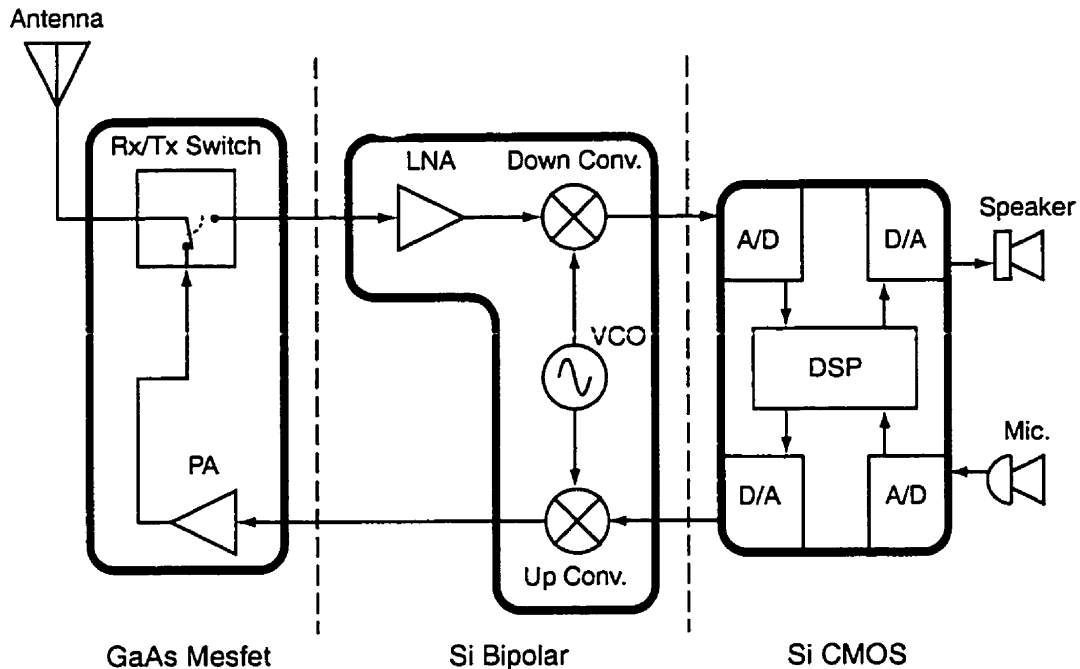


Figure 1.1: Typical Present-day Transceiver Configuration

1.1. Conventional Modern Transceiver Configuration

Figure 1.1 shows the configuration of a typical present-day wireless transceiver. As indicated, the transceiver is a fragmented design. Typically, devices are implemented in technologies that offer the lowest cost possible to achieve their respective performance specifications. This format results in a higher cost and lower performance than would ultimately be achievable in a fully integrated design. Interconnections between the various dies, for instance, result in parasitic losses. This degrades both the efficiency and signal strength of the transceiver. Not shown in the figure are the numerous discrete passive devices, such as filters and biasing resistors, that also should be integrated to optimize cost and performance.

The various components of a transceiver are implemented in the technologies that are the lowest possible cost to achieve adequate performance. Unfortunately, a typical transceiver has many different performance specifications. This necessitates the use of a number of different technologies. Typical output power amplifiers and receive/transmit switches are fabricated using gallium arsenide devices [3]. Gallium arsenide is a good technology to use because of its relative high performance. Higher

carrier mobilities result in higher transition frequencies, f_T , than would otherwise be possible using semiconducting substrates. This ultimately leads to RF components with higher gains and efficiencies than would be otherwise achieved. Lower fabrication yields and higher processing costs make gallium arsenide more expensive when compared to higher volume technologies though.

Low noise amplifiers, used as the first step in a receive chain, voltage controlled oscillators, and Gilbert-cell mixers are typically implemented in silicon bipolar or heterojunction silicon-bipolar processes. Although device transition times are not as high as experienced with gallium arsenide, good noise characteristics make bipolar a good alternative for the components mentioned. Silicon bipolar has the advantage that it can be fabricated in conventional CMOS production facilities with the addition of a few processing steps.

CMOS is presently the overall winner when price is considered. The present proliferation of personal computers have resulted in immense economies of scale associated with CMOS fabrication. Key digital signal processors and interface circuits used in today's cellular phones are also mostly implemented in CMOS. In consumer applications, price is usually the deciding factor in any purchase. It would therefore be ideal to integrate all of components used in the construction of a cellular phone onto one die using only CMOS as the chosen technology [4]. Considering that current CMOS devices capable of transition frequencies above 13 GHz, single die integration may soon be possible.

One of the most important elements of a transceiver is the transmit power amplifier, which is required to generate large amounts of radiated power to establish a good communications link with a radio tower. Because of the power levels required, the power amplifier has the greatest effect on overall cellular-mobile battery life. Careful consideration is therefore necessary to maintain efficiency performance when making design compromises to achieve our integration goal. The design of such integratable amplifiers was seen as a challenging design task and was therefore chosen as the focus of this work.

1.2. Frequency Standards

In both Canada and the US, there are primarily two bands used for mass public cellular communications. These bands have been allocated by Industry Canada and the Federal Communications Commissions respectively for public use with portable cellular devices. These bands are defined by the following frequencies:

AMPS band	824-849 MHz for transmit
AMPS band	869-894 MHz for receive
PCS band	1850-1910 MHz for transmit and
PCS band	1930-1990 MHz for receive

It is important to note that the transmit frequency refers to the cellular mobile and not the cellular tower [5]. Since this thesis was focused on transmit power amplifiers for the PCS band, the frequency range of interest was therefore chosen to be 1850 to 1910 MHz.

AMPS is a first generation cellular phone standard. Mobile phone voice signals are transmitted via a frequency-modulated carrier. AMPS stands for advanced mobile phone system. AMPS was developed before key signal processing technologies were available and therefore uses a complete analog communications link.

The PCS band is a relatively recent frequency allocation. There are a number of significant advantages with PCS over the AMPS standard. Although the increased frequency bandwidth associated with PCS is significant, the application of new digital modulation techniques have greatly enhanced the communications link. Digital modulation allows for digital signal processing techniques to be used on the transmitted signal. Among other things, this allows for significant capacity increases over its counterpart analog wireless link. The use of error correction techniques also enhances the quality of the communications link. Since the transmitted signal is digital in nature, the wireless link can also be readily adapted to the transmission of data. Finally, relatively simple encryption protocols can be readily implemented to provide levels of user privacy previously unachievable.

1.3. Modulation Standards

There have been a number of different communications standards chosen for use in the new PCS band. Table 1.1 summarizes these various standards. At the forefront of these modulation standards is code division multiple access, CDMA. With growing popularity and widespread acceptance, there are intrinsic advantages to CDMA that make this the outstanding standard of choice for communications networks in the future.

CDMA is a digital form of spread spectrum communications. Spectral spreading is a technique originally developed during the Second World War to prevent communications signals from being jammed by the enemy. Typically, both jamming signals and modern day spurious emissions are narrow-band in nature. This noise immunity characteristic therefore makes spreading an ideal choice for rejecting emissions that would normally block a conventional communications link. When a

Table 1.1: Modulation Standards Used for PCS Band Communications

Standard	IS-98	IS-54	PDC	GSM	DECT	PHS
Multiple Access Method	CDMA/ FDM	TDMA/ FDM	TDMA/ FDM	TDMA/ FDMA	TDMA/ FDMA	TDMA/ FDMA
TX Max Power	200 mW	600 mW	-	1 W	250 mW	80 mW
Duplex Method	Frequency Division Duplex	Frequency Division Duplex	Frequency Division Duplex	Frequency Division Duplex	Time Division Duplex	Time Division Duplex
Full Duplex	Yes	No	No	No	-	-
Channel Spacing	1.25 MHz	30 kHz	25 kHz	200 kHz	1.73 MHz	300 kHz
Modulation	OQPSK/ QPSK	Pi/4 DQPSK	Pi/4 DQPSK	GMSK	FSK	Pi/4 QPSK
Bit Rate	1.23 Mb/s	48.6 kb/s	42 kb/s	271 kb/s	96 kb/s	72 kb/s

narrow-band jammer blocks a spread spectrum signal, the amount of signal received is usually sufficient to maintain a high-quality link.

With spread spectrum, a narrow-band signal is dispersed over a wide frequency band. With the IS-98 standard, for instance, the required spectrum for single user could be as little as 9.6 kHz before spreading. After spreading though, this signal will be dispersed over a 1.2 MHz block of spectrum giving it immunity to narrow-band emissions. Fortunately, this immunity also allows for a number of other users to occupy the same block of spectrum at the same time. Along with noise and user immunity there is also a capacity increase over what would otherwise be achievable with AMPS or other frequency-division formats.

As the name implies, CDMA allows the simultaneous use of a block of spectrum by a number of different users. This is accomplished by a series of Walsh Codes. A typical Walsh Code is a binary sequence, 64 bits in length. Each of these Walsh Codes is orthogonal to each other meaning that each of the codes is uncorrelated. It is this characteristic of the zero cross-correlation that allows for a number of users to occupy the same channel at any point in time.

1.4. Transmit Power Standards

The IEEE and the American National Standards Institute have coordinated to generate a common standard for CDMA communications. IS-98 and J-STD-018 are the standards that have been generated by these respective standards agencies. Among other things, these standards specify the linearity and dynamic power ranges for the transmitter mobile. Class III for AMPS communication and Class II for the PCS have been defined to be in the range of -50 to 23 dBm. Spurious output levels are defined to be either, -42 dBc at an offset of 30 kHz, -60 dBm and an additional offset of 30 kHz, or -55 dBm for the field. The lower limit on transmission power is typically set by the signal strength needed for a mobile to reach a communications tower

1.5. Power Amplifier Options

One of the key design parameters in a power amplifier is efficiency. The following subsections briefly explain the key features associated with the amplifier classes most commonly used in high efficiency RF communications.

1.5.1. Class B

Class B amplifiers are more efficient than ordinary Class A configurations. The general configuration for this class is seen in Figure 1.2. Typical to this format is the inductive loading. The advantage of using an inductor over a resistor is that the DC voltage drop across an inductor is zero.

A 78.5 percent power-added efficiency is the maximum attainable for class B. This class is often used in high power applications where efficiency is important. In Class B, a typical device only conducts current for half of the conduction cycle. As seen in Figure 1.3, power is only be dissipated when the device is on. Since power dissipated in the transistor is voltage times current, a net power savings is realized.

In this amplification class, two devices can be used for pseudo-linear amplification. This configuration resembles the push-pull output stages used in audio applications. For Class B operation, each transistor is driven 180° out of phase through a necessary coupling transformer. Since there are no DC bias currents, no power is dissipated when there is no input signal.

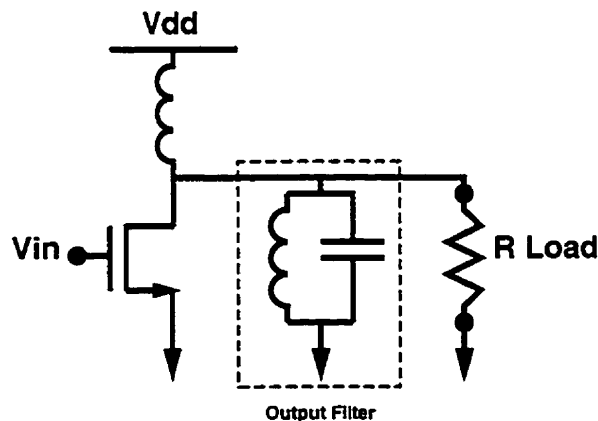


Figure 1.2: Generic Configuration for Class B, C, and E Amplifiers

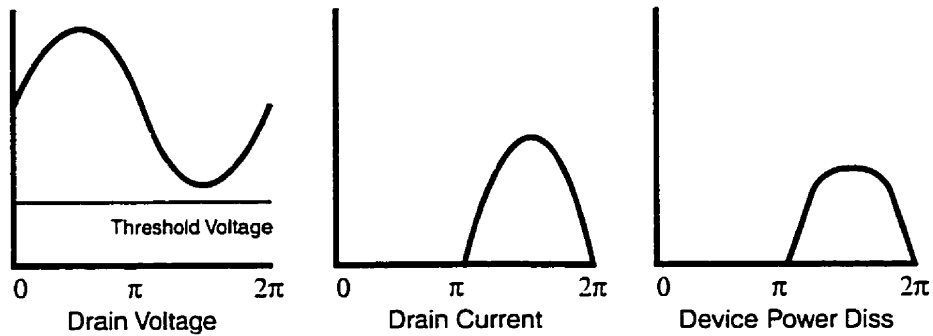


Figure 1.3: Class-B Waveforms

Crossover distortion can occur when the two amplifiers change phase. This reduces the overall linearity of the configuration. If there is an impedance mismatch between complementary PMOS and NMOS devices, a conduction imbalance can occur and ultimately result in additional harmonic distortion. This approach is ultimately desirable in applications where linearity is not critical.

1.5.2. Class C

Since not all applications require linear amplification, Class C was devised for further increases in efficiency. Radio stations usually use Class C in their transmitters where large transmit powers necessitate an efficient operating mode. As seen in Figure 1.4, the conduction angle for Class C is even less than that of Class B.

Unlike in Class A or Class B circuit designs, Class C has very high harmonic levels relative to the fundamental frequency. In this case, output matching network design is important to lower these unwanted signals to acceptable levels.

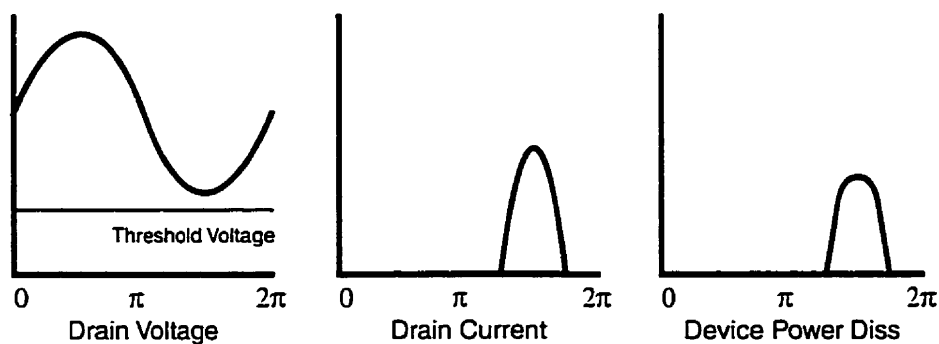


Figure 1.4: Class-C Waveforms

1.5.3. Class D

In a typical Class D amplifier, a pair of transistors act as switches to generate a square wave output voltage. The usual architecture that is considered to be Class D is shown in Figure 1.5. The two transistors, in this case, act in a fashion not unlike a CMOS inverter. The tuned output network ensures that only the fundamental frequency component is passed to the load. In theory, this mode of operation can achieve 100 percent efficiency. Finite transistor on-resistances and transistor switching times quickly degrade the performance of these amplifiers though.

In this amplifier format, drain capacitances aren't a part of the tuned output network and result in these capacitances being charged and discharged through the finite on-resistance of each transistor. Energy is therefore dissipated whenever these drain capacitances are charged and discharged.

Further losses occur if there is any mismatch between the two devices used. The resultant asymmetric switching can cause both transistors to be turned on at the same time. This can drastically reduce efficiency and potentially damage the amplifier's output devices.

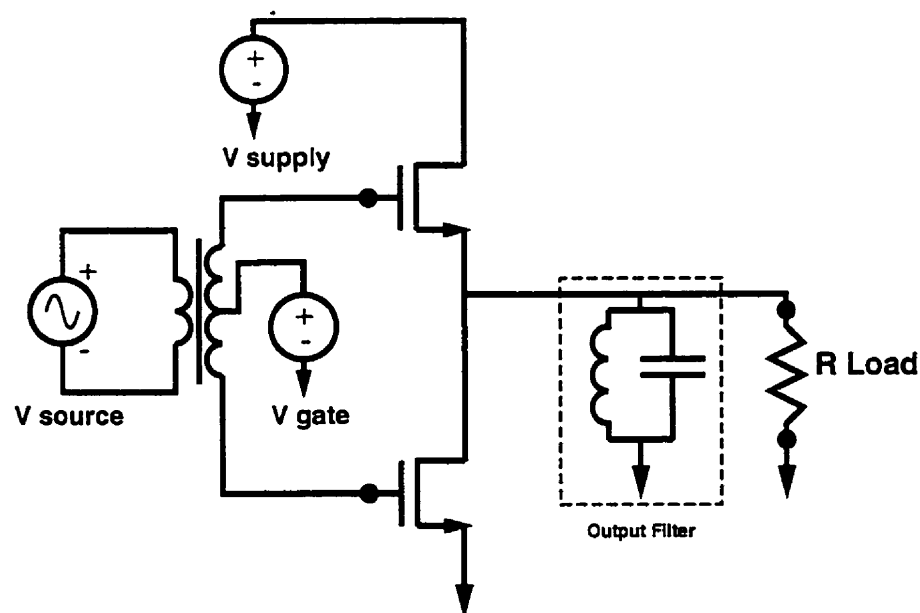


Figure 1.5: Class-D Amplifier

1.5.4. Class E

Class E is a relatively recent amplifier format. In this instance, as in Class D, the device is operated as a switch. In either case, any parasitic capacitance is usually detrimental. Unlike Class D though, the parasitic drain depletion-region capacitance can be resonated with the inductive load when configured as shown in Figure 1.2. Because of this, Class E can, in practicality, achieve very high efficiencies. Also contrary to Class D, there are no short-circuited currents that result in efficiency losses. This design also requires careful selection of the typical shunt-resonant circuit to reduce the high harmonic levels.

Figure 1.6 is an example of the typical drain voltage and current waveforms for this class. In this instance, the unique characteristic is that the drain voltage is driven well into the triode region. This key fact has significant positive effects on efficiency measures. Because the switch only turns on when the drain to source voltage is almost zero, there is very little power dissipation in the device. Power dissipation in the device is simply the drain-to-source voltage multiplied by the drain current. Therefore, when the drain voltage is reduced before drain current is drawn, efficiency increases. In fact, this form of amplification can ideally achieve 100% efficiency [6].

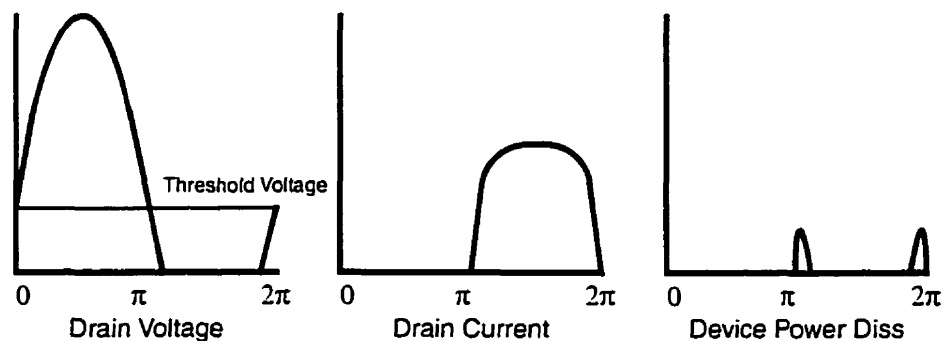


Figure 1.6: Class-E Amplifier Waveforms

1.5.5. Class F

Class F amplification was one of the first techniques developed for improving RF amplifier efficiency. This was first described by [7]. The general configuration for

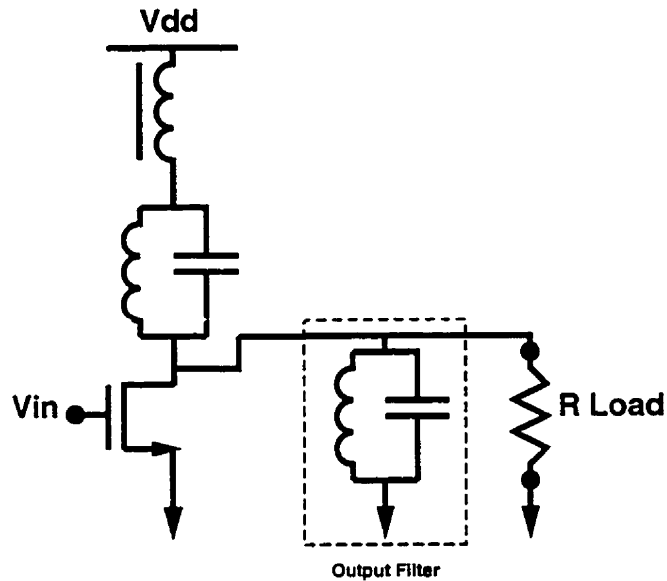


Figure 1.7: Class-F Amplifier Design

this amplifier can be seen in Figure 1.7. Class F uses an additional resonant network placed in the bias circuit to improve overall efficiency. This network is resonant at the third harmonic of the fundamental frequency. This results in an increase in the third harmonic component in the drain voltage waveform. As seen in Figure 1.8, a flatter waveform is evident when the transistor approaches the triode region. A net reduction in average drain-to-source voltage when the transistor turns on has positive effects on operating efficiency. This is very similar in operation to a Class E amplifier.

The transistor in a Class F amplifier acts as a current source similar to a Class B amplifier. Output filter design, in this case, is critical to prevent the peaked harmonics from reaching the load.

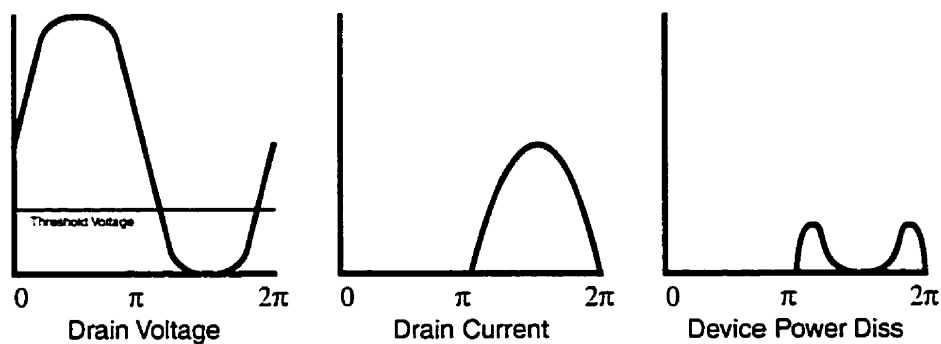


Figure 1.8: Waveforms for Class-F Amplification

1.5.6. Optimum Choice for a Power Amplifier Architecture

As seen above, there are a number of options available for a highly efficient amplifier class. Like Class D, Class E has the potential of achieving 100% efficiency. Since Class E uses only one device though, there is no chance of shoot-through currents to decrease efficiency. As well, Class E is a simpler amplifier to implement than Class F, where the design of a third harmonic resonator can be very difficult. Class E can therefore be seen as the optimum choice for an amplifier topology when all of these issues are considered.

1.6. Previous Work

There have been a number of attempts to implement high efficiency switchmode amplifiers using baseband CMOS technologies. The most notable of these contributions are shown in Table 1.2. Sowlati et al [6] used a 0.8 μm GaAs MESFET technology to achieve an efficiency of 57% with an output power of 200 mW. A production process which included on-chip inductors that were easily realized and a low resistance gold metallization used for interconnections contributed to the overall performance of the circuit.

Su and McFarland [8] used a 0.8 μm CMOS process for their implementation to successfully prove that CMOS could be used in RF applications to achieve good performance figures. Implementing in 0.8 μm CMOS most likely set the upper limit for their amplifier to the 825 MHz transmit frequency.

Tsai and Gray [9] devised a differential mode-locked configuration implemented in a CMOS process. This architecture produced a relatively high efficiency of 48 % and generated 1 Watt at 2 GHz. While the mode locked operation produces high efficiency and power gains, it has a number of disadvantages. Mode-locking, raises the lower limit on the input power necessary to maintain a frequency lock. This restricts the use of automatic gain control amplifiers in the transmit chain to perform envelope or power modulation required in CDMA applications. As also indicated in Tsai's work, mode locking causes the amplifier to oscillate even in the absence of an input signal.

This required an elaborate scheme to compensate for this problem. Finally, Tsai's design made use of only off chip matching networks to minimize input and output reflections.

Doyama [10] reported a single-ended design implemented in a 0.35 μm CMOS process. Although this amplifier was fabricated with an on-chip input matching network, the design was limited to the lower AMPS frequency.

Table 1.2: Previous Work on Class-E Power Amplifiers for Wireless Transceivers

	Sowlati, et. al [6]	Su, and McFarland [8]	Tsai, and Gray [9]	Doyama [10]
Technology	0.8 μm GaAs MESFET	0.8 μm CMOS	0.35 μm CMOS	0.35 μm CMOS
Frequency	1.9 GHz	824 - 829 MHz	2 GHz	835 MHz
Supply Voltage	2.4 V	2.5 V	2 V	3.3 V
Efficiency	57 %	42 %	48 %	36 %
Output Power	200 mW	1000 mW	1000 mW	79 mW
Die Area	5.7 mm ²	1.5 mm ²	0.8 mm ²	1.9 mm ²
Architecture	single ended	single ended	differential	single ended

1.7. Objectives and Outline of the Thesis

The main objective of this thesis was the design of a high efficiency power amplifier for use in PCS CDMA applications that is easily integrated with baseband digital circuitry. Since the overwhelming majority of baseband digital electronics is implemented in CMOS technology, and considering that CMOS with f_T values above

10 GHz are now available, a 0.35 μm CMOS process was chosen for the design of this amplifier.

Based on the previous discussion, the primary design criteria for this amplifier were determined to be, a target operating frequency of 1.9 GHz, optimal efficiency and a 200 mW power output. Differential operation and on-chip input matching network were identified as important features towards getting the amplifier closer to the integration goal. The supply voltage was chosen to be a standard 2.5 V. A class E configuration was also chosen for the design.

Chapter 2 contains a discussion of the design of a 200 mW CMOS amplifier. Chapter 3 includes an explanation of the layout considerations necessary for fabrication. Chapter 4 presents the experimental results and discusses the testing of the amplifier. This chapter will also present the full-power performance figures achieved by the final design. Chapter 5 presents the conclusions and recommendations for future areas of investigation.

1.8. References

- [1] J. Rapeli, *IC Solutions for Mobile Telephones*. Kluwer, 1993.
- [2] A. Matsuzawa, "Low-voltage and low-power circuit design for mixed analog/digital systems in portable equipment." *IEEE Journal of Solid-State Circuits*, vol. 29, pp. 470-480, 1994.
- [3] D. Gradinaru, "High Voltage RF Silicon Bipolar Transistor." MAsc. Thesis. University of Toronto, 1998.
- [4] J. Sevenhans and D. Rabaey, "The challenges for analogue circuit design in mobile radio VLSI chips." *Microwave Engineering Europe*, pp. 53-59, May 1993.
- [5] F. Ali, C. Wakeham, M. Williams, R. Moazzam and T.L. Ful, "RF Design Challenges for CDMA Cellular and PCS Mobile Handsets," *IEEE Radio Frequency Integrated Circuits Symposium, Proceedings*, pp. 7-10, 1998.
- [6] T. Sowlati, Y. Greshishchev, and C.A.T. Salama "1.8 GHz Class E Power Amplifier for Wireless Communications." *Electronic Letters*, vol. 32, pp. 1846-1848, 1996.
- [7] V.J. Tyler, "A New High Efficiency High Power Amplifier." *Marconi Review*, vol. 21, pp. 96-109, 1958.
- [8] D. Su and W. McFarland, "A 2.5 V, 1 W Monolithic CMOS RF Power Amplifier." *IEEE Custom Integrated Circuits Conference, Proceedings*, 1997.
- [9] K.Tsai, and P.Gray, "A 1.9 GHz, 1-W CMOS Class-E Power Amplifier for Wireless Communications." *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 962-970, 1999.
- [10] J. Doyama, "CMOS Class E Power Amplifier." M.A.Sc. Thesis. University of Toronto, 1999.

CHAPTER 2

Class-E Power Amplifier Design

2.1 Introduction

Sokal [1] first introduced the highly efficient Class-E amplifier in 1975. Through his research, Sokal demonstrated that this operating mode was capable of achieving nearly 100-percent efficiency. Later, Raab[2] realized the benefits of the Sokal's innovation and therefore expanded the work by using Fourier analysis to analytically describe the amplifier's characteristics. Zulinski furthered Raab's investigation through a series of works [3, 6]. In the end Zulinski derived an exhaustive set of equations that completely describe the Class-E frequency-domain condition.

Zulinski's complete solution is a system of over two hundred equations. Li[7] realized the design difficulties associated with such a high level of complexity, and therefore moved towards a refined approach. Li also devised techniques for reducing the size of the choke that Raab and Zulinski proposed. This would later prove to be beneficial in integrated applications.

Figure 2.1 is a diagram of the Class E resonant amplifier as summarized by Li. The choke inductor L , along with the drain capacitance C , are resonant close to the fundamental frequency. The transistor has traditionally been approximated as an ideal switch. L_r and C_r are series-resonant in order to block unwanted harmonics from

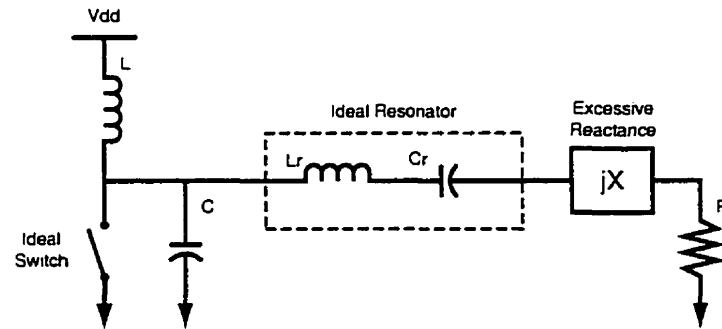


Figure 2.1: Class-E Resonator [1,7]

reaching the load. L_r and C_r are modeled as ideal components. Nonidealities associated with implementing the series resonator are lumped into jX , which is termed the excessive reactance. jX primarily serves to adjust the phasing in the L-C harmonic resonator. Voltage and current phase adjustments are used to maximize efficiency.

There are a number of approximations throughout the preceding derivations that permit analytic solutions. When operation with a CMOS amplifier at 2 GHz is attempted though, these approximations result in large errors. Firstly, for verification Zulinski used a frequency that was over three orders of magnitude lower than the intended frequency of this thesis. The parasitic interactions at 2 GHz would quickly invalidate the derived results. Modeling the transistor as an ideal switch also may be practical at 1 MHz but finite transition times and even the limitations of the most advanced transistor models themselves result in inaccuracies at 2 GHz [8]. Because of the bilateral nature of a MOSFET the feedback circuit must also be considered.

Another approximation of the mentioned analysis is the use of a 27 V power-supply. With the typical 2.5 V limit on modern CMOS devices, I^2R losses will be much more severe for a give output power. The assumption of high-Q discrete inductors is also unrealistic given the lossy inductors attained by CMOS integration.

Finally, lumped-element analyses also begin to break down at 2 GHz. The 5 cm wavelength experienced [9] sets the limit where distributed analysis becomes necessary at the chip level [10]. Complicated effects like fringing fields and skin-effect must also be considered as frequencies increase.

Although Li attempted to simplify the exhaustive Fourier analysis, his approach still resulted in convergence problems when calculations were attempted. The Sokal-Li

Fourier analysis was therefore concluded to far too complex for limited accuracy achieved. Although much of the previous work on integrated Class-E amplifiers was based on the Fourier-domain analytic solutions [11, 13], a much simpler time-domain approximation was devised and therefore adopted in its place.

2.2 Class E Theory of Operation

The analysis of the drain resonant circuit can be described by a simple set of differential equations, the solution to which is presented here. A simplified circuit model of a Class-E RF power amplifier is shown in Figure 2.2. This model initially replaces the transistor with an ideal switch. Later on, a more practical on-resistance was considered. When the switch is closed, current flows from the DC power supply through the inductor, L , and finally through the switch to ground. When the switch is opened, energy stored in the inductor causes current to continue flowing through it. Since current can no longer flow through the opened switch, current is diverted into either the drain capacitance, C , or the output load R . When the above process is repeated at the carrier frequency, current that is forced into the resistive load each cycle

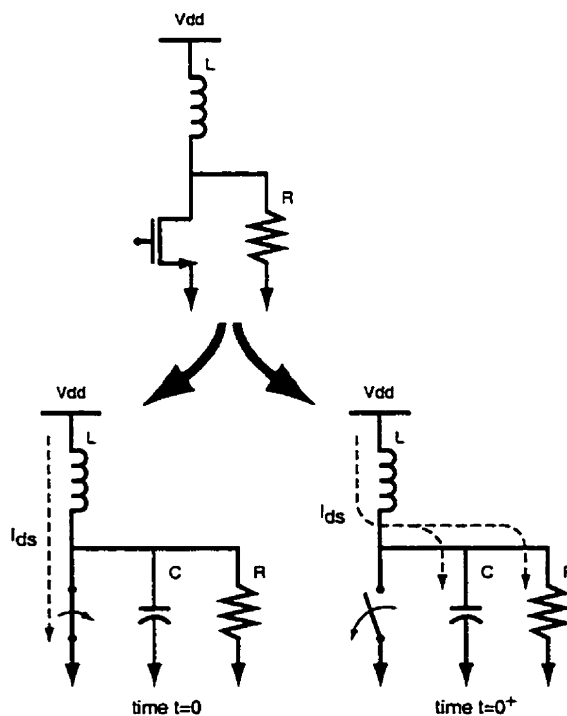


Figure 2.2: Simplified Drain Resonant Circuit

becomes the output power generated by the amplifier.

This model obviously ignores higher-order parasitic elements and focuses only on the first-order components that have a primary effect on oscillation. In reality, a complete empirical model of the output stage would consist of dozens of discrete devices. A number of these components become significant at RF frequencies. Although a first order model compromises accuracy in the result, a simplified approach is necessary to allow an intuitive understanding of the circuit. Once a general circuit has been determined though, the level complexity and accuracy can be greatly increased with the prudent use of a circuit simulator.

In addition to understanding the parasitics, timing for the drain resonant circuit is also very important. Figure 2.3 illustrates the problems that can occur when this circuit is improperly designed. In the center two frames, we notice the waveforms for a properly timed Class E amplifier. Both the voltage and current waveforms in this case are non-overlapping. The device is timed in such a way that the switch closes and begins to conduct current exactly when the drain-to-source voltage reaches a minimum. Power dissipated as heat in the output device is therefore minimized when the product of the drain-to-source voltage and drain current is decreased.

For sub-optimally tuned Class-E amplifiers, both power and efficiency performance degrade rapidly. In the case where the drain resonant frequency is too

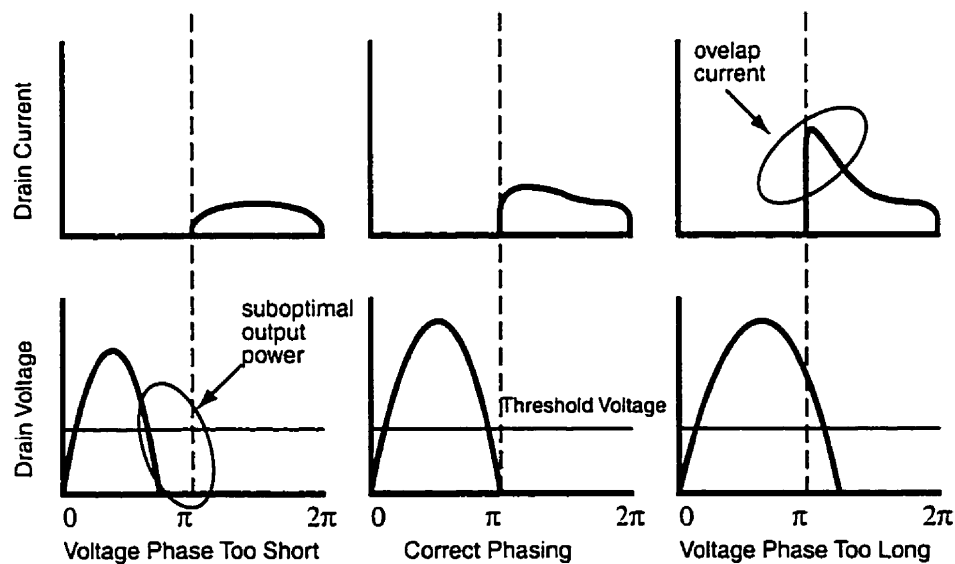


Figure 2.3: Drain Voltage Waveforms

high, there will be a reduction in output power from the ideal value. Improper tuning causes a reduced voltage swing from what would otherwise be possible. Since a smaller sine wave is presented to the antenna load, this results in decreased output power. Although not indicated in the diagram, a potential problem with this operating state is the occurrence of a negative drain-to-source voltage. In this situation, the reverse-biased drain can lead to permanent damage. The underdamped nature of this operating mode can additionally lead to oscillation.

In the case where the resonant frequency of the drain network is too low, efficiency falls off drastically. Overlapping voltage and current waveforms mean that a significant amount of power will be wasted in the output device. In this case, the residual drain voltage causes additional current to be driven into the output device when the transistor turns on, compounding the power dissipation problem. Since power amplifiers typically deal with large currents and voltages, thermal dissipation problems can also lead to device failure.

Under closer examination of the circuit in Figure 2.2, it becomes apparent that operation of this amplifier can be described as a simple parallel RLC resonator. Figure 2.4 is a representation of this circuit immediately after the switch is opened. At this instant, the current flowing through the inductor is diverted into either the resistor or the capacitor.

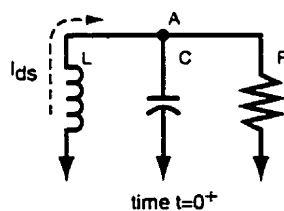


Figure 2.4: Simplified Drain Resonant Circuit

A homogeneous linear differential equation can be devised to describe the time domain behavior of the resonant system. According to Kirchhoff's current law, the current entering node A can be described by the following equation:

$$C \cdot \frac{dv}{dt} + \frac{v}{R} + \left(\frac{1}{L} \cdot \int_0^t v(u) du - I_o \right) = 0 \quad (2.1)$$

Differentiating the above yields the following result:

$$\frac{d^2 v}{dt^2} + \frac{1}{R \cdot C} \cdot \frac{dv}{dt} + \frac{1}{L \cdot C} \cdot v = 0 \quad (2.2)$$

The time domain solution for the voltage in the differential equation can be solved using the quadratic equation:

$$\lambda_1 = \frac{-\frac{1}{R \cdot C} + \sqrt{\left(\frac{1}{R \cdot C}\right)^2 - \frac{4}{L \cdot C}}}{2} \quad \lambda_2 = \frac{-\frac{1}{R \cdot C} - \sqrt{\left(\frac{1}{R \cdot C}\right)^2 - \frac{4}{L \cdot C}}}{2} \quad (2.3)$$

The above coefficients are used in the following time domain solution:

$$v(t) = D1 \cdot e^{\lambda_1 \cdot t} + D2 \cdot e^{\lambda_2 \cdot t} \quad (2.4)$$

For the final solution to the differential equation, one needs to solve for the constant coefficients. Constant coefficients, D1 and D2, should therefore be determined from the initial conditions of the system. The time domain equations for the system at time $t=0_+$ take the following form:

$$v(0_+) = D1 \cdot e^{\lambda_1 \cdot 0_+} + D2 \cdot e^{\lambda_2 \cdot 0_+} \quad (2.5)$$

The initial conditions for the voltage level in the system is determined by the charge on the capacitor:

$$v(0_+) = \frac{Q_0}{C} \quad (2.6)$$

Differentiating the time domain solution yields the following equation at time $t=0_+$:

$$\frac{dv(0_+)}{dt} = D1 \cdot \lambda_1 \cdot e^{\lambda_1 \cdot 0_+} + D2 \cdot \lambda_2 \cdot e^{\lambda_2 \cdot 0_+} \quad (2.7)$$

The initial condition for the differentiated equation can now be found with the following:

$$C \cdot \left. \frac{dv}{dt} \right|_{0^+} = - \left(\frac{Q_0}{C} \cdot \frac{1}{R} + I_0 \right) \quad (2.8)$$

As an approximation, the capacitor can be assumed to be fully discharged at time $t=0$. This is in fact a reasonable assumption based on the previous discussion. The constant coefficients for the time domain equation will now have the following basic relationship:

$$D1 + D2 = 0 \quad (2.9)$$

Using the same approximation of a fully discharged capacitor, the differentiated time domain solution now has the following relationship:

$$\left. \frac{dv}{dt} \right|_{0^+} = D1 \cdot \lambda_1 \cdot e^{0^+} + D2 \cdot \lambda_2 \cdot e^{0^+} = -\frac{I_0}{C} \quad (2.10)$$

Simplifying the above equation:

$$D1 \cdot \lambda_1 + D2 \cdot \lambda_2 = -\frac{I_0}{C} \quad (2.11)$$

Therefore one can now solve for both constant coefficients, D1 and D2, and the exponential coefficients λ_1 and λ_2 for a complete solution to the expression.

Once these time domain equations are solved, two possible approaches can be taken to calculate the optimal values of the elements in this resonant network. The first approach entails solving the above equations for a periodic frequency at the carrier rate. This approach is similar to the one chosen by Zulinski et al. As well as being more complicated, this technique yields little intuitive information on the selection of device parameters. The second approach considered, and eventually applied, was an iterative solution of the above equations for the RLC circuit parameters. Since the system of equations chosen to define the circuit's operation is relatively easy to solve, numeric solution is now straightforward. Plotting iterative results allowed for a visual confirmation of correct operation and the quick convergence to a solution.

2.3 Class-E Amplifier Design

The actual construction of the output stage begins with the power requirements for the overall design. This parameter was defined by the standards for the communication protocol discussed in Chapter 1. For Class II PCS CDMA communication, the total power requirement is 200 mW at the fundamental. Since this project is a differential design, 100 mW is required from each half-circuit.

In the previous section, our Class-E amplifier was broken down into a simple three-element resonator. To define these three elements, the analysis must begin with load resistance. Since the power requirement for each half of the amplifier is known, voltage swing is subsequently required to determine the load resistance. Drain-to-source breakdown voltage sets the upper limit on the signal-swing. For the 0.35 μm technology used in this design, a reasonable drain-to-gate breakdown voltage was found to be of 5 volts. Higher voltages would risk the possibility of junction breakdown and permanent device damage.

The optimal value for the load resistance is determined by means of a load-line analysis. To determine the value of the output resistance, the following equation can be used. For a peak-to-peak swing of 4.5 volts, the value of this resistance was determined to be:

$$R_{load} = \frac{\left(\frac{V_{pk-pk}}{2\sqrt{2}}\right)^2}{P_{out}} = \frac{\left(\frac{4.5}{2\sqrt{2}}\right)^2}{0.1} = 25.3 \quad (2.12)$$

Therefore, a 25-ohm load impedance is required for each half of the differential circuit for optimal loading. This value will ultimately maximize the power output from the amplifier.

There are a number of criteria to consider when specifying the transistor used in the output stage. In addition to the timing considerations discussed previously, device on-resistance plays a major role in the efficiency performance of a Class-E amplifier. On resistance determines the amount of I^2R power dissipated during the device's conduction cycle. Also, the associated drain-to-source voltage-drop reduces the amount current available to charge the drain-loading inductor, reducing overall RF power generated. Channel length is a design variable that can control the amount of resistance

encountered. Channel length should be kept to a minimum in applications such as this where switching-speed limits circuit performance. Since the current conduction primarily occurs in the triode region, channel length has the following relationship with on-resistance [14]:

$$r_{ds} = \frac{l}{\mu_n \cdot C_{ox} \cdot \left(\frac{W}{L}\right) \cdot V_{eff}} \quad (2.13)$$

μ_n represents mobility, C_{ox} is oxide capacitance, W , L , and V_{eff} are channel width, length, and effective-voltage respectively. A minimal channel length will therefore reduce the drain resistance value. In addition, gate-to-bulk capacitance is also reduced preventing unnecessary loading of the associated driver stage. For this project, the minimum channel length technology available at the time of project's implementation was $0.35 \mu\text{m}$.

Since drain-to-source resistance is so important in determining amplifier performance, after examining equation 2.13 it becomes apparent that further benefits can be gained by making the channel width large. The compromise with arbitrarily increasing channel width is that peripheral drain-to-bulk capacitance will also increase linearly. This relationship is evident in the following definition of this parameter:

$$C_{db} = A_d \cdot C_{jd} + P_d \cdot C_{j-sw} \quad (2.14)$$

A_d and P_d represent drain area and perimeter respectively. C_{jd} and C_{j-sw} are defined as the respective drain-junction and sidewall capacitances.

Excessive drain capacitance is an undesirable consequence of using a large channel width to reduce on-resistance. This is evident because drain capacitance restricts the natural frequency of the time-domain solutions derived in the previous section. This relationship can also be shown through the equation for the resonant frequency of a parallel tank-circuit:

$$f = \frac{1}{2 \cdot \pi \cdot \sqrt{L \cdot C}} \quad (2.15)$$

In this application, a reduced natural frequency will eventually lead to voltage and current phase overlap, and as a result, a reduced operating efficiency described earlier.

Before determining the resonator circuit values, the element that defines the initial condition must be identified. Since transistor dimensions can be made arbitrarily large or small, the transistor therefore does not restrict the subsequent result. Drain inductance though, is limited by the dimension of the smallest achievable bondwire used to connect the circuit for a minimum value, and limited by series resistance for a practical maximum value. Referring to equation 2.15 again, it is apparent that reducing the resonator's inductance will have a positive effect on performance by increasing the drain resonant frequency. Decreasing the size of the drain inductor will also permit a larger device to be used for a given frequency of interest. Drain inductance can therefore be used to define the allowable capacitance for a given output device and hence the size of the device itself.

The smallest value achievable for a drain inductor was found to be 2 nH. This value was derived from the geometry of the shortest possible bondwire used for drain loading. The relationship between the drain capacitance and device on-resistance was also needed in order to model the drain resonance. On resistance and inductor series resistance were both later used in computations to characterize inductor current at time $t=0_+$.

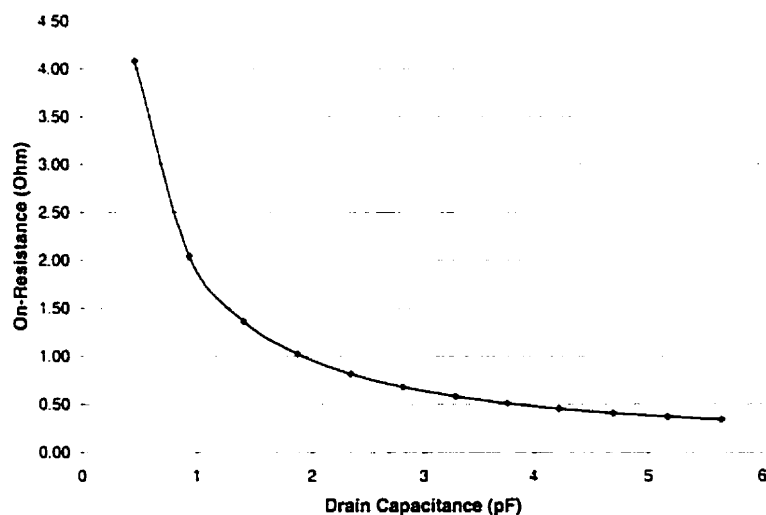


Figure 2.5: Drain Capacitance and On-Resistance Proportionality

Transistor models supplied by the fabrication facility were used to define the necessary transistor variables. Initially, the proportionality of drain capacitance to device on-resistance was extracted from simulations. This relationship is shown in Figure 2.5. Once this proportionality was known, work could proceed with solving for the optimal resonant condition.

The defined parameter values were input into the system of equations developed in the previous section. These equations were solved iteratively for varying values of transistor parameters while maintaining inductance constant. Plotting these results allowed for the quick identification of the optimal Class-E condition. The result of one of these sweeps is shown in Figure 2.6. Seen in this diagram is a plot of the drain resonance versus swept drain capacitance and associated on resistance. Plotting the

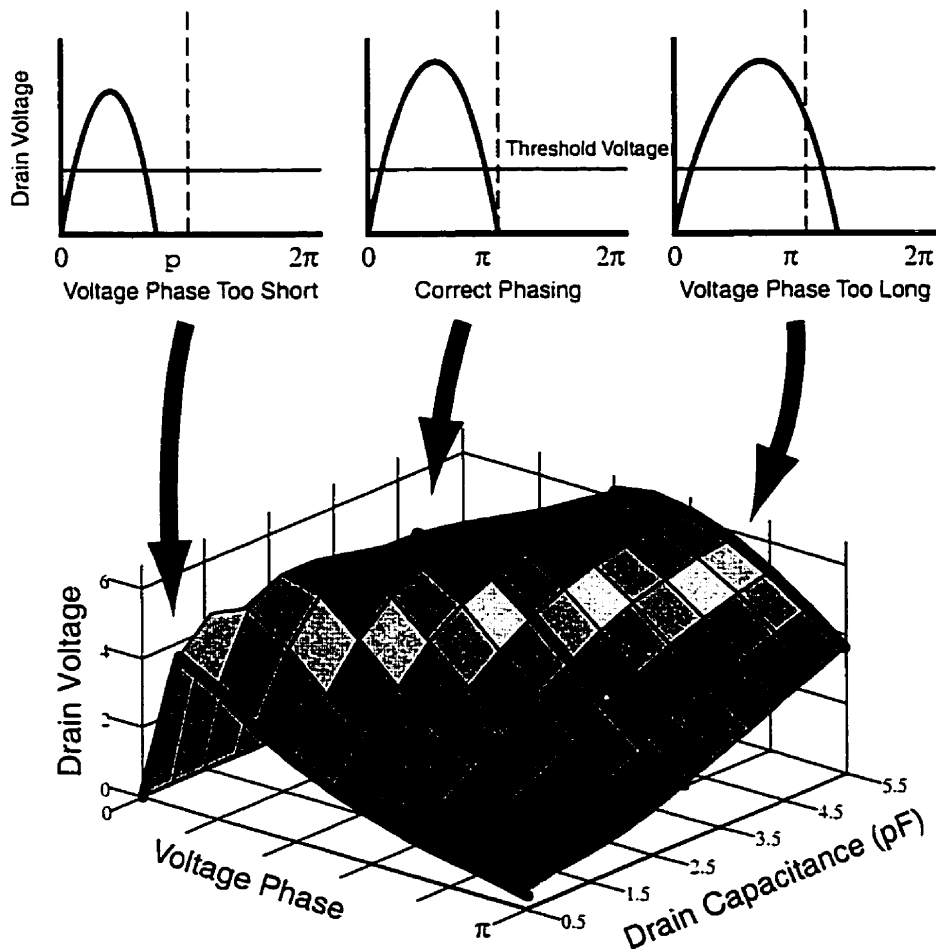


Figure 2.6: Drain Resonant Waveform

data allowed for the quick determination of the best variables to use in order to produce the optimal drain resonant waveform. The optimal drain capacitance illustrated in Figure 2.6 was found to be approximately 2.75 pF.

Knowing the required drain capacitance permitted the final selection of transistor channel width. The relationship of the channel width to the drain capacitance was used to make this determination. A plot of this relationship can be seen in Figure 2.7. As seen in the chart, a capacitance of 2.75 pF translates into a channel width of 3.0 mm. This was therefore the channel width selected for the output device.

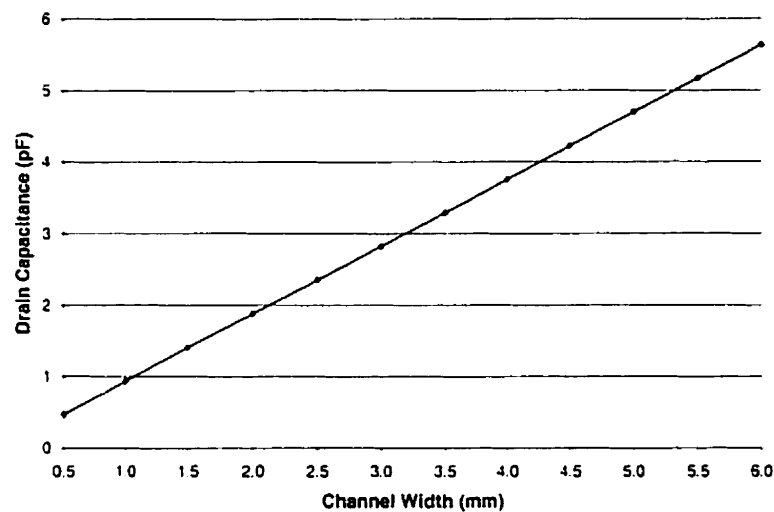


Figure 2.7: Relationship of Drain Capacitance to Channel Width

Figure 2.8 is a final simulation plot for output stage functioning at the desired carrier frequency for one half of the differential amplifier. Drain voltage and source current waveforms show little overlap in order to optimize efficiency. Drain voltage waveform shows no evidence of potentially damaging negative voltage. A residual drain-to-source voltage can be seen in the voltage waveform. This is a result of the finite device on-resistance. The drain voltage waveform can also be seen to be free of any undesirable oscillations that are possible at high operating frequencies.

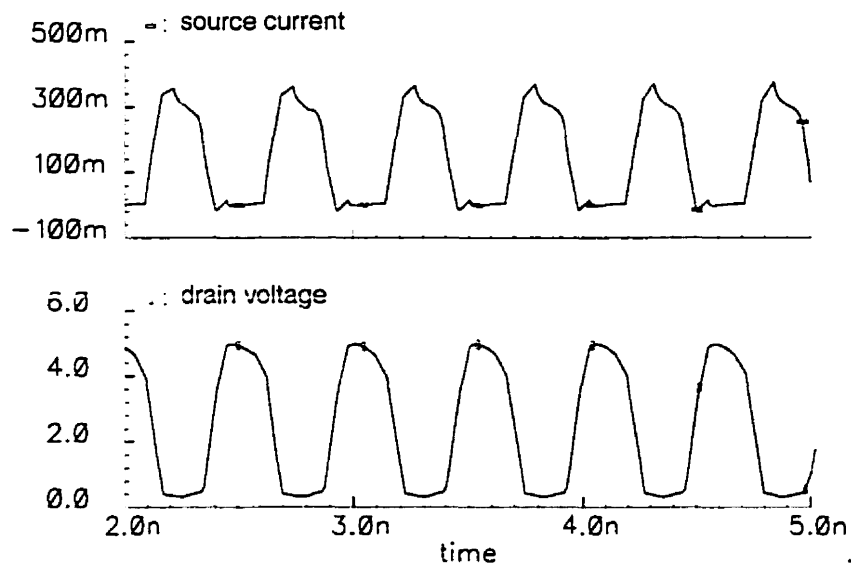


Figure 2.8: Final Amplifier Waveforms

2.4 Differential Architecture

There are a number of valid reasons for choosing a differential architecture over a single ended design. Firstly a differential architecture serves to protect highly sensitive low noise amplifiers and receive circuits operating at or near the transmit frequency. Receive circuits are designed to amplify extremely small radio signals and are therefore very sensitive. As a result, these circuits are susceptible to the high substrate current levels generated by power amplifiers. Figure 2.9 is an indication of one potential feedback path followed by these currents.

A differential amplifier has beneficial isolation properties because the amplifier sources current into the substrate at twice the carrier frequency. Because each half of a differential circuit can be driven at opposite phase, current is sourced into the substrate each half cycle. This is equivalent to current being sourced at twice the carrier rate. In single-ended designs, substrate current simply presents itself at the transmit frequency [11]. Figure 2.10 is an illustration of how this process occurs. When a layout is carefully generated with particular attention paid to symmetry and spacing, cancellation benefits can readily be realized.

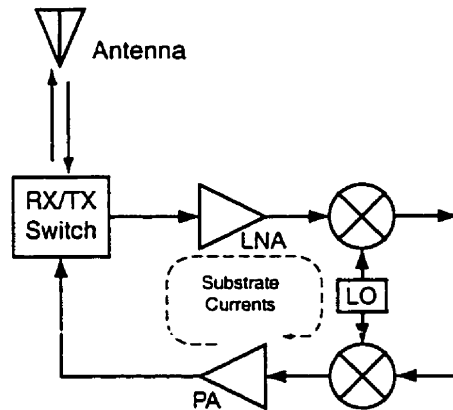


Figure 2.9: Substrate Current Effects on Receiver Circuit[15]

There are power benefits that can also be gained by using differential construction. As suggested in Figure 2.10, a differential topology generates twice the signal swing of a single-ended format. This is an important benefit with the low operating voltages associated with emerging technologies. A lower operating voltage means a necessary reduction in load resistance to generate the same amount of output power. Since load resistance value can rapidly approach the resistive losses seen in the interconnects, the benefits associated with an increased differential swing voltage are vital.

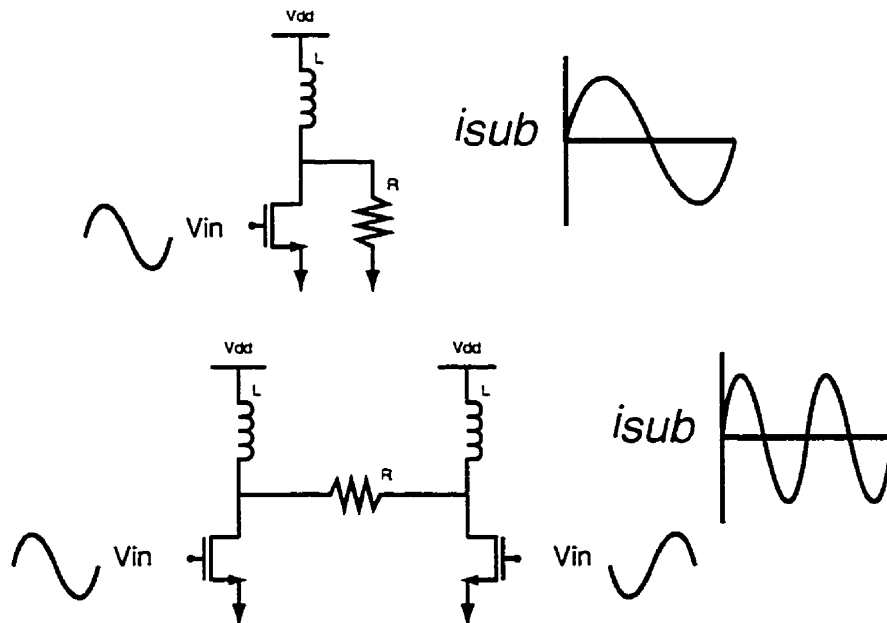


Figure 2.10: Comparison of Single-Ended and Differential Substrate Currents

When the input leads to a differential amplifier are closely spaced in a wire-bonded implementation, the fields associated with the input leads cancel and reduce the associated lead inductance over what would otherwise be observed. This is opposite to the effect experienced in coiled inductors where the mutual inductance of current-carrying conductors running in parallel combines to produce a larger overall inductance.

Figure 2.11 gives an indication of how cancellation was accomplished in the implementation. As indicated, input bondpads were placed next to each other. Input leads were run as close together as possible thereby reducing lead inductance.

Since the end goal of this implementation was a fully integrated design, it was desired to integrate the input matching network on-chip. The use of a differential topology provided an additional benefit in this regard. Since there are two input matching networks in parallel instead of just the one network which would be used for a single ended amplifier, matching resistive losses are halved when compared to a single-ended counterpart.

Finally, the choice of a differential topology makes it convenient to drive the amplifier with a Gilbert-cell mixer. Gilbert cell mixers produce a differential output that would otherwise require the addition of a differential to single-ended transformer when used with a single ended amp.

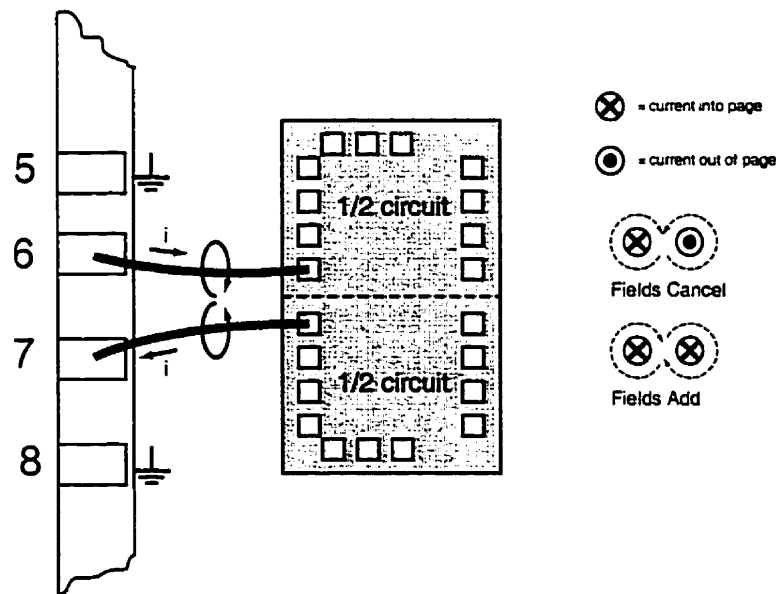


Figure 2.11: Mutual Inductance Cancellation

2.5 Bondwire Inductors

The performance of the drain inductor is one of the key aspects of a Class E amplifier. The impedance of this component drastically effects the amplifier's natural frequency, damping factor, and ultimately the overall power and efficiency.

Unfortunately, high quality-factor integrated inductors in standard CMOS technology are very difficult to achieve. Recently there has been a considerable amount of activity in this field to improve the capability of these devices. Typical present day planar spiral inductors as big as 10 nH can be implemented in silicon semiconducting substrates. The maximum frequency of these devices is limited to only a few gigahertz though [16].

Advanced processing techniques can be used to achieve high quality-factors with spiral inductors. The use of on-chip gold metallization, insulating substrates, air bridges, and thick polyamide depositions used for the inductor substrate can all be used to increase the performance of these devices [17]. Etching techniques offer the possibility of greatly increasing inductor self-resonant frequencies. This technique allows for inductors as large as 100 nH to be integrated on-chip for frequencies in the low-gigahertz range [18].

Attempts to reduce the series resistance of these devices through the use of wide metal traces and multiple metal layers does, in fact, lead to a lower series resistance but can also result in additional substrate capacitance and a corresponding lower self-resonant frequency

An alternative to the use of on-chip spiral inductors is to take advantage of the parasitic inductance associated with a chip's interconnection bondwires. The integration of these components into the actual circuit has a number of advantages. The first advantage of utilizing these components is that there is a net die-area reduction when compared to a similar circuit implemented with on-chip spirals. Since a typical on-chip spiral drain inductor would consume most of the die area in an RF amplifier circuit, the space savings associated with utilizing bondwire inductance is considerable.

An equally important benefit to using bondwire inductors is the increased quality factor, or Q-factor. When gold is used to make an interconnection, resistance is minimized. Since the geometry of a bondwire can be controlled, and also considering

that the distance between a bondwire and any conducting surface is typically large, parasitic capacitance can be seen to become negligible when compared to a spiral.

Bondwire inductors were also chosen in this design instead of spiral inductors because of low resistive losses and a low noise contribution to the circuit. In a typical CMOS process, using the top metal layers can reduce capacitance to the substrate. A typical 10 nH inductor, for instance, can therefore operate up to a gigahertz. The series resistance for such a device, though, remains at about 15 ohms. Lower series resistance can be obtained by using wide metal paths and several metal layers in parallel with multi-layer process but, once again, the losses in the substrate are unavoidable in standard silicon processes.

The geometry of an actual bondwire can vary depending on chip orientation and the fixture used for testing. Because of this, it can be difficult to predict bondwire characteristics a priori. Parameter values can be extracted by experimentation though. In a production implementation, the accuracy of modern manufacturing equipment should be sufficient to give confidence in the extracted values.

2.6 Source Impedances

A serious consequence of high frequency design is that the series impedance of the package leads and associated bondwires becomes large at gigahertz frequencies. The series inductance of a lead in a common high performance package can be as high as three nanohenries. A general rule of thumb RF designers use in amplifier design is one nanohenry per every millimeter of bondwire length. This means that a relatively short 3 mm bondwire will have a primary impedance of 3 nanohenries associated with it. At 2 GHz, the three-nanohenry inductor translates into a 38-Ohm imaginary impedance. An impedance this large will greatly attenuate RF signals when placed in a signal path. In the input and output signal path, bondwire inductance can be addressed by resonating out this impedance at the carrier frequency. In the amplifier's source and drain, though, the need to carry dc currents and the likelihood of instability prevents the use of series-capacitive resonant techniques.

Another technique commonly used to compensate for bondwire inductance is the application of on-chip supply decoupling capacitors. Although this technique is

very effective in mitigating the effects of the drain and source bondwires, both drain and source impedances will be affected. Since this design is heavily dependent on bondwire inductance for drain loading, the cancellation effects experienced by supply decoupling would be highly undesirable.

Fortunately, steps can be taken to minimize the effects of packaging. Since Ohm's law applies to two impedances run in parallel as it does for two resistances, multiple bondwires can be used in parallel to reduce the overall impedance of a particular signal path. This approach is identical to the one used in the microprocessor industry to reduce the amount of series resistance seen in the ground path of a chip. Since typical microprocessors operate at much lower frequencies, and given that the large packages typically employed afford a great number of additional power leads, RF packaging considerations can therefore be more challenging in comparison.

Figure 2.12 is a diagram of the leads used in the test fixture to define the circuit ground plane. This layout made use of the maximum number of leads possible to define the circuit's ground path. Unfortunately the large number of leads needed for other circuit functions limited the number of leads available for the ground path. To a

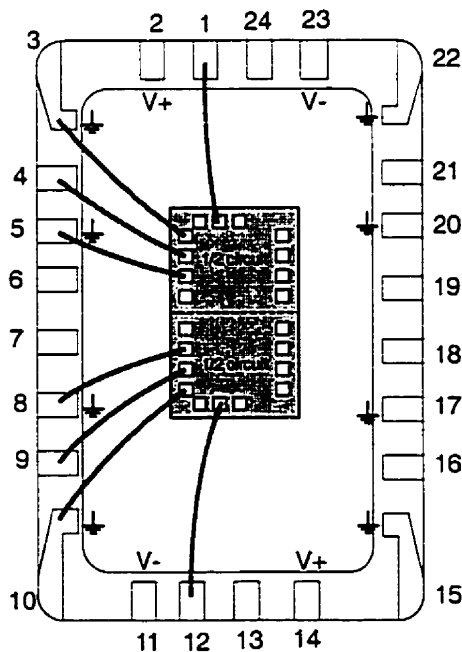


Figure 2.12: Test Fixture Ground Plane Configuration

significant extent, the configuration of the ground plane determines the general layout of the amplifier. Prior knowledge of this parameter is therefore essential for a successful implementation.

Figure 2.13 shows the model for the resultant ground plane for one half of the differential amplifier. The limited number of leads available meant that grounding presented to the amplifier was far from ideal. This has the effect of raising the voltage level present at the source. The term applied to this phenomenon is groundbounce. The net effect of groundbounce on amplifier performance is a reduced gate to source voltage available to drive the amplifier. Ultimately, this leads to less power gain and a reduced gain compression point.

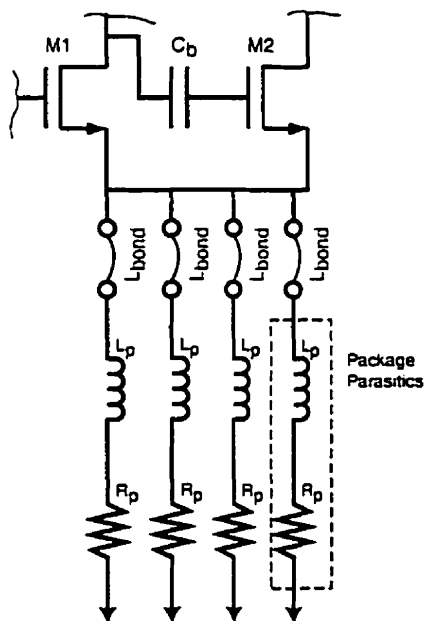


Figure 2.13: Ground Plane For Half Circuit

The effects of grounding non-idealities are illustrated in Figure 2.14. Shown in this figure is the gate voltage applied to the half-circuit. Also seen in the illustration is the resultant ground level seen at the source for the full half-circuit using the actual ground-plane impedance experienced by the final implemented amplifier. It is apparent that the gate-to-source drive is drastically reduced. This obviously reduces the net power available to drive the amplifier and the resultant power output.

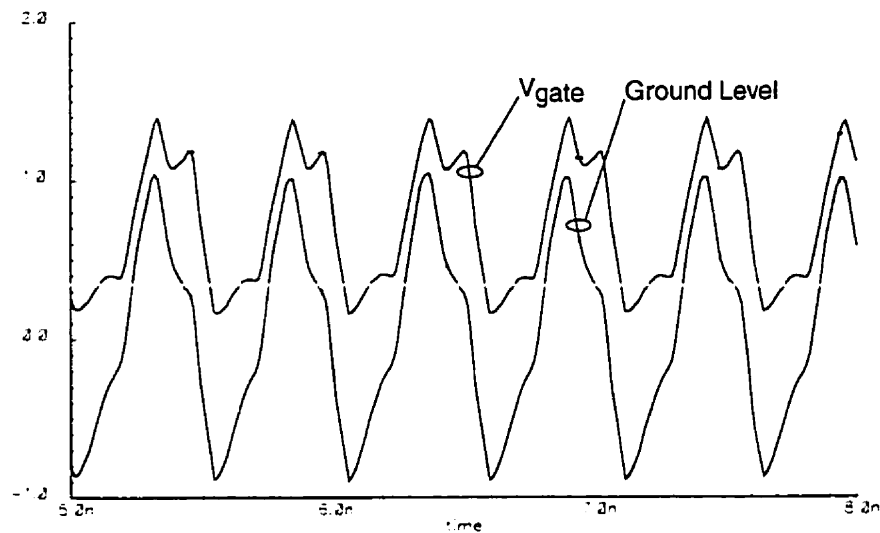


Figure 2.14: Effects of Non-Ideal Grounding

2.7 Gain Stage Design

Ideally, a class Class-E power amplifier would be driven with a square wave input. The circuit best suited to accomplish this task would be a CMOS inverter. Unfortunately, at two gigahertz, a CMOS inverter implemented in $0.35\ \mu\text{m}$ technology would be very inefficient. The charging and discharging of large depletion region capacitances each cycle would result in large amounts of power being wasted.

A circuit better suited to the task of producing a signal close to a square wave at 2 GHz would be a power amplifier operating in Class-F mode. Mentioned briefly in Chapter 1, Class-F operation is defined by an additional resonance at the third harmonic of the fundamental frequency. Harmonic peaking is typically accomplished with the addition of a tank circuit placed in the drain of the amplifier. Because this circuit must be resonant at the third harmonic, this translates into a frequency of 5.7 GHz for a 1.9 GHz fundamental. The difficulties associated with accurately designing an efficient tank circuit resonant at such a high frequency with conventional semiconducting substrates and uncharacterized on-chip spiral inductors were deemed to be too great a risk to the overall project. Also, the large amount of die area required for two

additional spirals to achieve marginal gains in efficiency was seen as another disadvantage of this approach.

A Class-E amp would have been another suitable candidate for the gain stage. Although Class E amplifiers are highly efficient, they typically must operate well into gain compression in order to achieve their characteristic performance. Although the high efficiencies are desirable, efficiency isn't as important in the initial gain stage where less power is required. An amplifier with a higher gain and similar efficiency performance was therefore more desirable. The inability to apply any type of load tuning, or load pulling, on an interstage device was another detriment to using Class-E. Since the drain resonant waveform could not be probed or tuned after the implementation as with the output stage, any miscalculation on a level of parasitics would have meant that the circuit would not operate properly in Class-E mode.

Among the amplifier candidates introduced in Chapter 1, both Classes B and C are the best suited to driving a Class-E output stage. There are a number of benefits in choosing non-linear operating Class B or C amplifiers, these include simplicity of design, a small chip area consumed, and a relatively high operating efficiency. Since a Class B is rigidly defined to have a 50% conduction angle, Class C was the chosen format for the gain stage to allow for phasing latitude after the implementation.

The analysis of a Class C amplifier can be approximated by the equations used to describe a Class B design when the Class C amplifier is operated close to a 50% conduction angle. Since Class-B design is a relatively straightforward process, the equations derived by Sedra [19] can be used as the basis for an amplifier design.

The device used in a Class-C amplifier can be modeled as a current source and not like a switch as in the Class-E case. Current draw, in this case, is simply determined by device saturation current and not on-resistance.

Current in the amplifier can be approximated by the following equations:

$$i_{drain} = i_b \cdot \sin(\omega_0 \cdot t), \quad i_{drain} > 0 \quad (2.16)$$

$$0, \quad i_{drain} < 0 \quad (2.17)$$

These equations represent the familiar truncated sinusoidal current expected with Class-B operation, where i_b is the magnitude of the truncated sinusoid. Since the majority of

RF power is resident in the fundamental frequency, for ease of computation, only the fundamental needs to be considered. Calculation of the magnitude of the fundamental component of current is determined by integrating over the amp's duty-cycle:

$$i_{\omega_0} = \frac{2}{T} \cdot \int_0^{T/2} i_b \cdot \sin(\omega_0 \cdot t) \cdot \sin(\omega_0 \cdot t) \cdot dt = \frac{i_b}{2} \quad (2.18)$$

Drain voltage can simply be determined by multiplying the magnitude of fundamental current, i_{ω_0} , by the impedance presented to the drain. Drain voltage in terms of fundamental and transistor current is therefore:

$$v_{drain} \cdot \sin(\omega_0 \cdot t) = Z_{drain} \cdot i_{\omega_0} \cdot \sin(\omega_0 \cdot t) = Z_{drain} \cdot \frac{i_b}{2} \cdot \sin(\omega_0 \cdot t) \quad (2.19)$$

Peak efficiency is achieved when the magnitude of the voltage swing is made as large as possible. Because of inductive biasing, voltage swing is centered upon the supply rail. The supply voltage can therefore be used to determine the maximum output power based on the following equation:

$$P = \left(\frac{Vdd}{\sqrt{2}} \right)^2 \cdot \frac{1}{Z_{drain}} = \frac{Vdd^2}{2} \cdot \frac{1}{Z_{drain}} \quad (2.20)$$

The DC input current is now determined from the device drain current.:

$$i_{ave} = \frac{1}{T} \cdot \int_0^{T/2} \frac{2 \cdot Vdd}{Z_{drain}} \cdot \sin(\omega_0 \cdot t) \cdot dt = \frac{2 \cdot Vdd}{\pi \cdot Z_{drain}} \quad (2.21)$$

Therefore, DC input power is simply defined by:

$$P_{in} = i_{ave} \cdot Vdd = \frac{2 \cdot Vdd}{\pi \cdot Z_{drain}} \cdot Vdd \quad (2.22)$$

The input power should obviously be more than the output power, as it is here, for an efficiency less than 100 %.

Efficiency can now be found by dividing output power by input power:

$$\eta = \frac{P}{P_{in}} = \frac{\pi}{4} \quad (2.23)$$

Determining the specifications for the transistor is now a relatively straightforward. A minimum channel length is desired in this instance based upon the

same assumptions used for the Class E amp. The device, in this case, operates as a current source. The saturation current equation can therefore be used to define the transistor:

$$i_b = \frac{k'}{2} \cdot \frac{W}{L} \cdot (V_{gs} - V_t)^2 \quad (2.24)$$

The Class B efficiency equations introduced above can define peak current, i_b . k' is a transistor characteristic parameter supplied by the fabrication facility and is therefore constant. The only variable required to determine channel width is the gate-to-source voltage. Gate-to-source voltage has a wide degree of freedom because it is controlled by the designer during device testing. The only restriction on this parameter is the device's gate breakdown voltage.

From the efficiency derivation, we know that transistor peak current is:

$$i_b = \frac{v_{drain}}{Z_{drain}} \cdot 2 = \frac{V_{dd}}{Z_{drain}} \cdot 2 = \frac{2.5}{144} \cdot 2 = 34.7 \text{ mA} \quad (2.25)$$

Where Z_{drain} , once again, is the impedance presented to the drain of the amplifier. Rearranging the saturation current equation, channel width can now be determined as follows:

$$W = \frac{I_b \cdot 2 \cdot L}{k' \cdot (V_{gs} - V_t)^2} = \frac{0.0347 \cdot 2 \cdot 0.35}{2.7 \times 10^{-5} \cdot (1.3 - 0.55)^2} \cong 1500 \mu\text{m} \quad (2.26)$$

The approximate value for the channel width required was therefore determined to be 1500 μm . This value was found for an approximate drive voltage of 1.3 volts.

2.8 Load Pull Analysis

Load-pull analysis is a technique with which an impedance match is optimized. A special technique is necessary because small signal S-parameters are not very practical for high power amplifier design. For instance, VRG's vector network analyzer generates a maximum output power of -10 dBm. As a point of comparison, the amplifier implemented in this thesis requires approximately 10 dBm of input drive at full power to function properly.

High-power amplifiers will seldom conjugate match for optimal power output. A wide variety of matching conditions are possible to tune a power amplifier though. Matching conditions for; optimal power, optimal efficiency, optimal noise, and optimal gain are all possible when tuning an RF amplifier [20]. Each of these matching states will typically be different from one-another. Optimal gain is the only amplifier condition that is generally understood to be attained with conjugate matches at both the input and output.

Load pull analysis is a replacement for the cut and try iterative approach used to devise an optimal matching network. The input match to an amplifier will be conjugate for an optimal power condition. The output match for optimal power, though, will not be a conjugate match. The gain achieved by matching for optimal power will also be less than the ideal gain value. This value should also be less than the Maximum Available and Maximum Stable Gain values when stability is considered.

Since gain and amplifier output impedance cannot be used as guidelines to determine a load impedance and with the mentioned difficulties associated with predicting and amplifier's implemented characteristics, one must iteratively sweep for the optimal impedance value. By varying both the magnitude and phase of the output load impedance one can plot the contours of constant power on a Smith Chart in order to determine the optimal load value. To generate these contours, the amplifier is biased at a preset current. The magnitude and phase of the load are then varied. Output power is measured and recorded on a Smith chart after each these variations. The result of this study will be a figure that resembles the plot in Figure 2.15. Each of the contours in the figure represents a constant output power. By observing the patterns generated, load impedances that produce a high output power can be identified [21].

Varying the DC bias and input signal levels, series of different contours can be identified. Although this would permit the more thorough discovery of the optimal biasing and matching condition, this approach would greatly increase the number of measurements required.

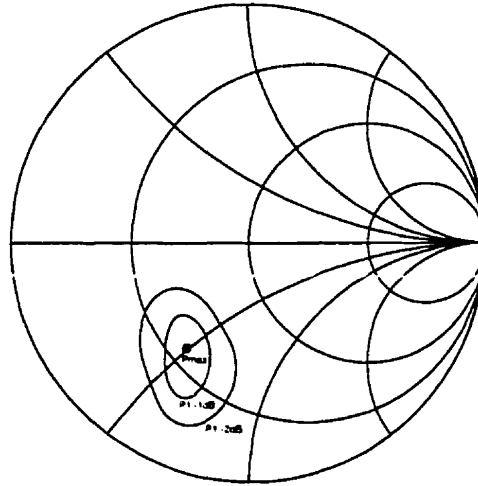


Figure 2.15: Typical Load Pull Power Contours

2.9 Impedance Matching

2.9.1 Spiral Inductor Design

The design of planar spiral inductors in semiconducting processes requires careful consideration. Even small miscalculations in the values of these elements when used for impedance matching can result in match that is far from what was desired. Unfortunately, most CMOS fabrication facilities do not supply a characterized set of inductor models as is typical with higher performance processes. Without the luxury of a number of iterations to extract parameter values, a designer is forced to rely on computational methods to determine these characteristics.

Wheeler [22] presented a simple formula that can be used to quickly estimate the approximate inductance of a square spiral inductor implemented on an arbitrary substrate:

$$L = \frac{45 \cdot \mu_0 \cdot n^2 \cdot a^2}{22 \cdot b - 14 \cdot a} = \frac{45 \cdot 4 \cdot \pi \cdot 10^{-7} \cdot n^2 \cdot a^2}{22 \cdot b - 14 \cdot a} \quad (2.27)$$

The letter a , in this case, represents the average radius of the windings in meters. This parameter is illustrated in Figure 2.16 clarification. Parameter b , accounts for the distance from the center of the windings to outer edge. The parameter n is used to account for the number windings in the implemented coil.

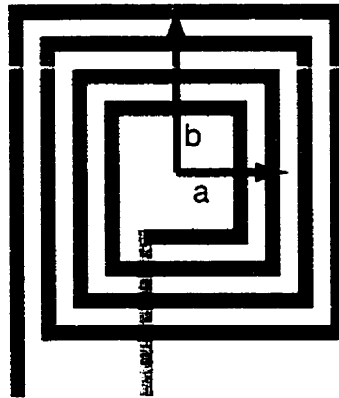


Figure 2.16: Illustration of Spiral Inductor Modeling Parameters

Unfortunately, Wheeler's equation lacks the detail necessary for multi-gigahertz design and should therefore only be used to give a first order estimate of inductance for use in hand calculations. Yue compiled a number of the higher-order parasitics into a physical model that can be used to expand Wheeler's calculation into one more appropriate for high frequency applications. In general, this model is referred to as the Pi model. Under single-ended excitation, as implemented in this design, the Pi model takes the form seen in Figure 2.17. The additional parameters introduced in the illustration are now defined as follows:

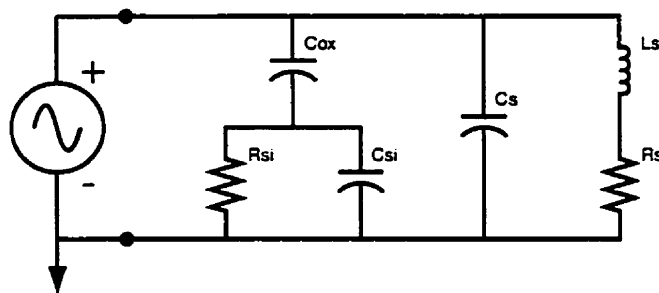


Figure 2.17: Single-Ended Pi Model for CMOS Spiral Inductor

$$\delta = \sqrt{\frac{2 \cdot \rho}{2 \cdot \pi \cdot f \cdot \mu_0}} \quad (2.28)$$

$$R_s = \frac{\rho \cdot l}{w \cdot \delta(\Gamma) \cdot \left(1 - e^{-\frac{l}{\delta(\Gamma)}}\right)} \quad (2.29)$$

$$C_s = n \cdot w^2 \cdot \frac{\epsilon_0 \cdot \epsilon_r}{t_{\text{mlm}2}} \quad (2.30)$$

$$C_{\text{ox}} = \frac{1}{2} \cdot l \cdot w \cdot \frac{\epsilon_0 \cdot \epsilon_r}{t_{\text{ox}}} \quad (2.31)$$

$$C_{\text{si}} = \frac{1}{2} \cdot l \cdot w \cdot C_{\text{sub}} \quad (2.32)$$

$$R_{\text{si}} = \frac{2}{l \cdot w \cdot G_{\text{sub}}} \quad (2.33)$$

Where δ represents skin depth. R_s represents skin-depth adjusted series resistance. C_s is the capacitance of the under-pass. C_{ox} is the capacitance of the coil to the silicon substrate. C_{si} is the capacitance through the silicon substrate to the back-side ground plane and R_{si} represents the silicon substrate resistance. The constants used in the above equations are listed in Table 2.1. These parameter values are based on the physical constants of the 0.35 μm process.

Results from the above analysis were validated with Agilent Technologies' EEsof Momentum planar electromagnetic simulator. Momentum is a high-performance solver that uses the method-of-moments to determine S-parameters for distributed microwave devices. The device, as simulated, is shown in Figure 2.18.

Table 2.1: 0.35 μm CMOS Process Inductor Design Parameters

N	4.5 turns
A	60 μm
B	110 μm
ρ	0.032 Ω -
l	3155 μm
w	10 μm
ϵ_0	$4\pi \times 10^{-13}$
ϵ_r	3.9
t_{m1m2}	0.8 μm
t_{ox}	3.8 μm
C_{sub}	1.6×10^{-18}
G_{sub}	4×10^{-8}

Although the results produced by Momentum are highly accurate, computations for even relatively simple devices with high-end processors can take hours to complete. As an example, the final segmented planar spiral illustrated in Figure 2.18 required 56 minutes to solve using a RISC-based Sun Ultra-10 machine equipped with 500 megabytes of RAM. Obviously when a designer must do many iterations to determine the appropriate geometry for a matching condition, the delays associated with an

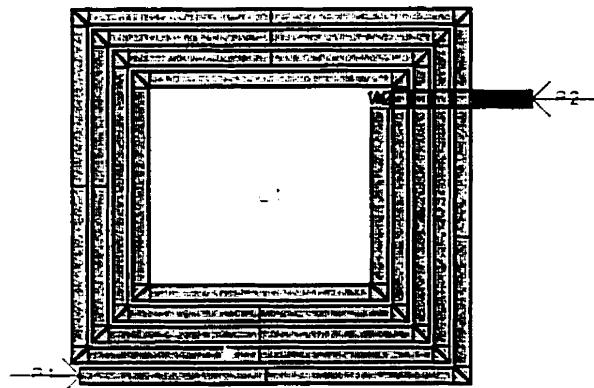


Figure 2.18: Spiral Inductor Layout as Simulated in EEsof

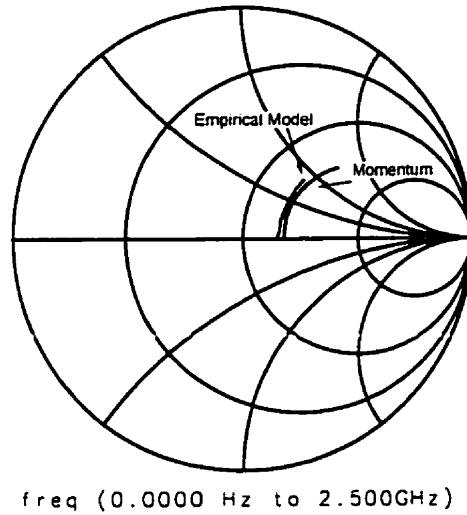


Figure 2.19: Comparison of Empirical Inductor Model and Electromagnetic Analysis

electromagnetic solver are unacceptable. Instead, an optimal process for designing inductors would be to use lumped-element models for iterative solutions and an electromagnetic solver as a validation technique.

A comparison of the reflection coefficients for the inductor calculated by the empirical method and calculated by the method-of-moments approach is shown in Figure 2.19. As the chart indicates, there is a significant difference in the inductive

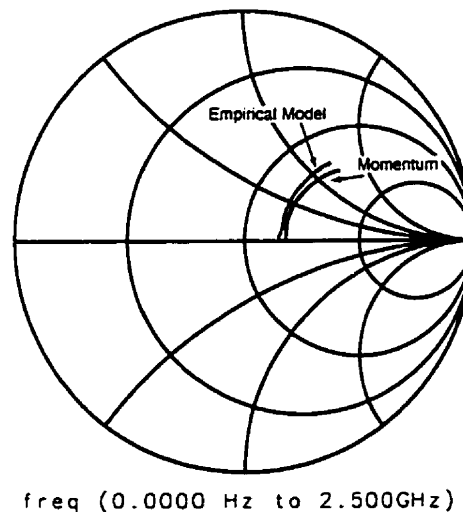


Figure 2.20: Comparison of Compensated Empirical and Electromagnetic Inductor Analysis

reactance calculated by the two approaches. Since the accuracy of the electromagnetic technique has been well documented, it was assumed that the physical model was in error. A closer examination of the physical model revealed an underestimate in the series inductance calculated by Wheeler's equation. This is understandable since effects like coil-to-coil coupling and substrate spacing are not accounted for in Wheeler's estimate. Increasing in the value of series inductance by 1.5 nH yielded excellent agreement between these two techniques. The final s-parameters for the corrected physical model and the electromagnetic analysis can be seen in Figure 2.20.

2.9.2 Input Matching

The input reflection coefficients for the amplifier are seen in Figure 2.21. This is a reflection plot for one half of the differential circuit. Two-element matching was used to conjugate match the input while the output of the amplifier had an optimized load presented to it. Using the inductor designed in the previous section, the amplifier was conjugate matched as shown in Figure 2.22. The orientation of the matching circuit can be seen in the final amplifier layout.

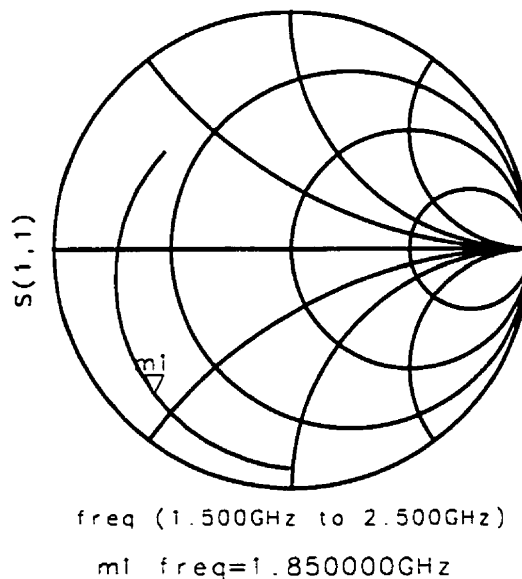


Figure 2.21: S-parameter Sweep of the Amplifier Input Reflection

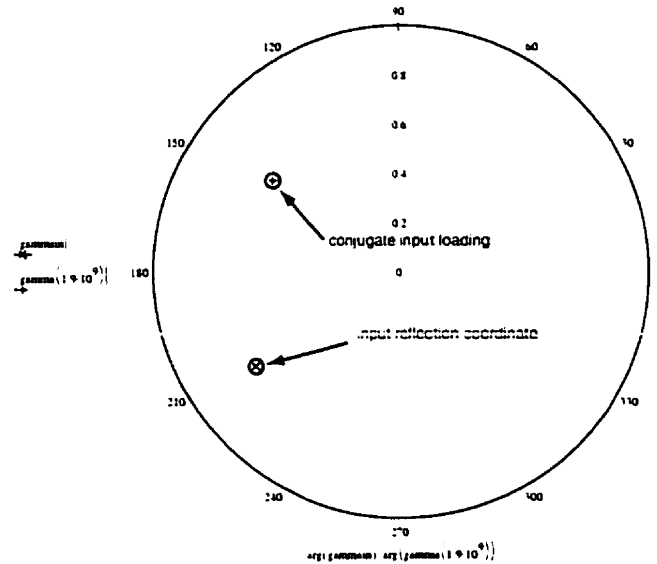


Figure 2.22: Input Reflection and Conjugate Input Match

2.9.3 Output Matching

The primary objective for the output-matching network was to use load pulling to discover the optimal load condition. Because of the need to tune the output stage after device fabrication, output matching was therefore implemented off-chip. It should also be noted that off-chip matching will be possible in a fully integrated CMOS transceiver.

2.10 Thermal Modeling

With this project being a high power implementation, special attention was paid to the effects of heat dissipated in the package. A good understanding of temperature effects is necessary to determine an amplifier operating point and to prevent the amplifier from ultimately being damaged. Fortunately, there are a number of useful texts on the subject, those by Holman [23] and Janna [24] are just two good examples of the many books that could provide an excellent reference.

A relatively simple model of the mounted chip in a high frequency Alumina package was constructed and is shown in Figure 2.23. This illustration is intended to be the side profile for the package introduced in Figure 2.12

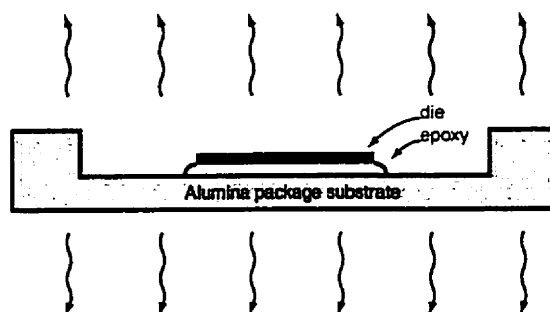


Figure 2.23: Simple Thermal Packaging Model

The calculations were broken down into two parts. First to be considered is the convection resistance portion, R_c . This heat transfer component encompasses thermal conduction to the air. This is similar to the heat transferred by touch but the medium in this case is air. Radiation resistance on the other hand, is similar to the infrared heating. In this case though, there is no medium involved in the transfer of heat. The symbol used for radiation resistance in these calculations is R_r .

Convection resistance for the top side plate was calculated to be:

$$R_{c_top} = \frac{1}{h_c \cdot A_s} = \frac{1}{(12.6 \cdot 10^{-3}) \cdot \left(\frac{0.75}{2.54} \cdot \frac{1}{2.54} \right)} = 682.709 \frac{^{\circ}\text{C}}{\text{watt}} \quad (2.34)$$

h_c is the top-side convection resistance and A_s represents the surface area.

Convection resistance for the bottom side plate is calculated as:

$$R_{c_bot} = \frac{1}{h_c \cdot A_s} = \frac{1}{(6.3 \cdot 10^{-3}) \cdot \left(\frac{0.75}{2.54} \cdot \frac{1}{2.54} \right)} = 1.3654 \cdot 10^3 \frac{^{\circ}\text{C}}{\text{watt}} \quad (2.35)$$

Bottom side convection resistance is slightly higher because of the poorer air circulation for the bottom plate.

The analogy between thermal resistance and electrical resistance holds. Total convection resistance is the parallel combination of the top and bottom plate resistances. This value is calculated as follows:

$$R_c = \frac{R_{c_top} \cdot R_{c_bot}}{R_{c_top} + R_{c_bot}} = 455.139 \frac{^{\circ}\text{C}}{\text{watt}} \quad (2.36)$$

The computation of radiation resistance is similar to that for convection resistance. In this case though, the energy radiated by the top and bottom plates is not affected by air circulation and will therefore be the same for each side. One equation can therefore be used to calculate the total radiation resistance:

$$R_r = \frac{1}{\tau \cdot h_r \cdot A_s} = \frac{1}{0.45 \cdot (4.5 \cdot 10^{-3}) \cdot \left(\frac{0.75}{2.54} \cdot \frac{1}{2.54} \right) \cdot 2} = 2.124 \cdot 10^3 \frac{^{\circ}\text{C}}{\text{watt}} \quad (2.37)$$

As indicated, radiation resistance plays a lesser role in determining the amount of heat transferred from an object than does convection resistance.

The final thermal resistance for the packaging is, once again, the parallel of the convection and radiation resistances:

$$\frac{R_r \cdot R_c}{R_r + R_c} = 374.821 \frac{^{\circ}\text{C}}{\text{watt}} \quad (2.38)$$

The above implies that a when 100 mW is dissipated by the amplifier into the substrate, this will result in a temperature rise of 37 degrees. In reality, the temperature rise will be much worse. For a 200 mW power output an approximate anticipated efficiency of 40 %, the computed radiation resistance predicts a worst case temperature rise of 187 degrees Celsius at the device junction.

A thermal resistance this high will obviously cause problems. Manufacturers will not recommend the designer implement designs that will allow junction temperatures to exceed 150⁰ C. To surpass the recommended temperatures will not only greatly hamper device performance, but will also increase the likelihood of device breakdown.

Since the limiting factor in this amplifier's thermal design was found to be the package's convection resistance. A chip-on-board mounting technique was necessary to solve this problem. This technique, described later in the thesis, uses a very large heatsink and would therefore allow the thermal dissipation to approach the near ideal

values calculated by Hashjani [25]. Before a chip-on-board technique could be implemented though, the package described above was used for low power validation testing.

2.11 Final Amplifier Schematic

Figure 2.24 is the final amplifier schematic as integrated on-chip. As seen in the illustration, two identical, non-coupled amplifiers are driven out of phase to form a differential amplifier. Drain loading is accomplished via hi-Q bondwire inductors. Included in the model are the packaging parasitics, because they have a significant effect on the circuit's performance. Input matching inductors are represented as simple inductors for clarity.

Input drive method and output loading are also indicated in the illustration. The input is driven with a 180-degree hybrid. Inputs to each half of the differential amplifier are fed 180 degrees out of phase by means of this device. Although the power splitters are lossy their effects were deducted from the overall performance of the circuit. In a typical implementation, the inputs to the circuit will be driven differentially by a Gilbert Cell mixer. In this case the outputs of a Gilbert cell do not require modification to drive this design. A differentially driven antenna can replace the power splitter in the output path. This configuration would resemble the one indicated previously in Figure 2.10. For developmental purposes though, the use of an antenna was highly impractical. For single-ended antennas, a balun should be used to transform the differential signal.

Output matching is performed off-chip to minimize resistive losses. This is the best design choice since power must be transferred off chip to drive an antenna. In this case, the use of high Q surface mount components is a far better choice than integrating in the lossy environment on chip.

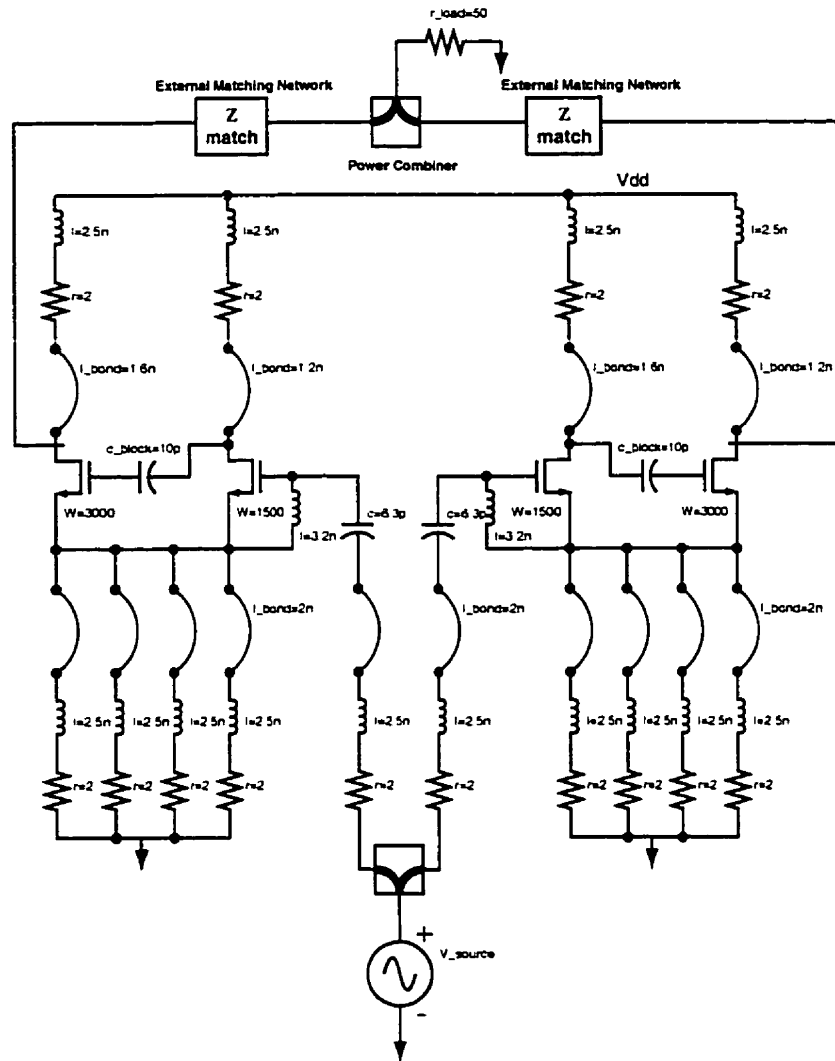


Figure 2.24: Final Amplifier Schematic

2.12 Theoretical Performance

Figure 2.25 is a simulation plot for swept load reflection magnitude and phase. Load magnitude and phase are simply the load S11 reflection coefficients for the load as presented to the amplifier. The reflection coefficient for any particular load can be computed by the following relationship:

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (2.39)$$

Z_L is the complex value of the load impedance and Z_0 is the reference impedance. The value typically used in RF engineering is 50 Ohms. Reflection coefficients can be any complex number with magnitudes ranging from 1 to -1 , these correspond to load magnitudes ranging from infinity to zero respectively.

Examining the figure once again, one can see that the phase of the load impedance is labeled on the horizontal axis while the numerous iterations are for varying load magnitudes. Power is labeled on the vertical axis. An output power level of 23 dBm translates into the power goal of 200 mW.

It should be noted that the accuracy of the values calculated by this approach have a significant potential for variation. Since output matching is performed off-chip, bondwire lengths and test fixture characteristics will play a major role in determining the actual amplifier output impedance and necessary load impedance. Although the

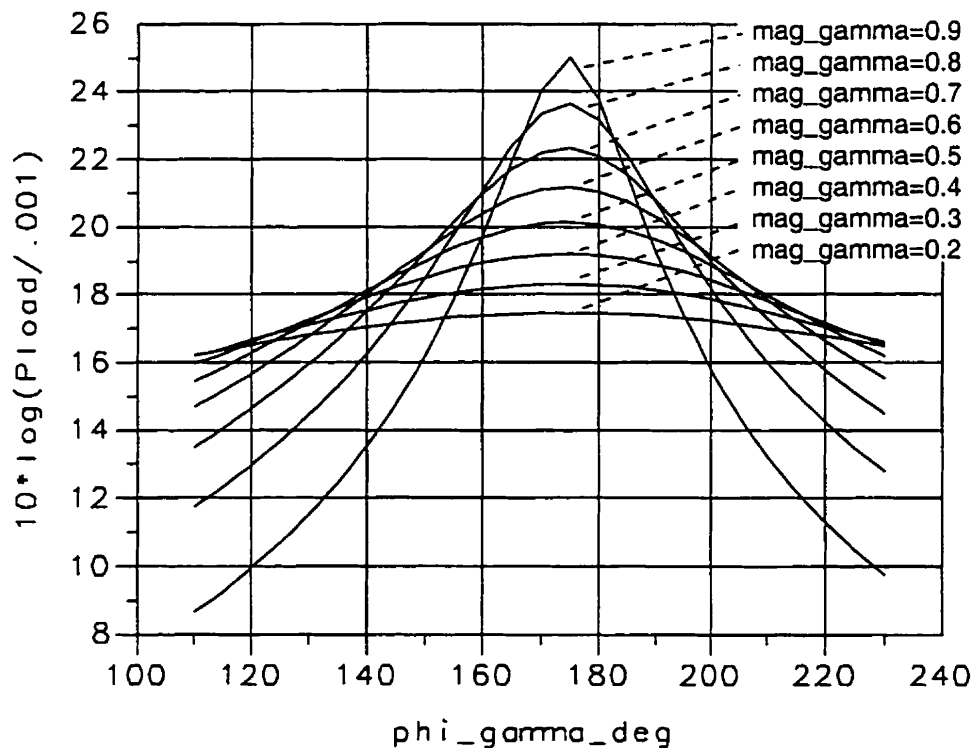


Figure 2.25: Simulations for Varying Loads Reflections for the Differential Amplifier

difficulties associated with impedance estimation have been thoroughly described, the plot should give an indication of the sensitivity of the amplifier to impedance variations.

Figure 2.26 is a plot of the predicted output power and efficiency for the full differential amplifier. The plot is generated for an output impedance optimized using the technique described in the load-pull discussion. As seen in this case, 3 dB gain flatness extended over the desired 1.85 to the 1.91 GHz bandwidth. Gain rolloff is a function of the single-frequency two element L-C matching. Broadband matching could have been used, and would have greatly simplified the implementation. A broadband match would have resulted in much worse performance figures than those presented here though. Efficiency performance is seen to follow power output as expected.

Figure 2.27 displays the amplifier's theoretical performance versus supply voltage. Power output and efficiency are both affected by this parameter's variation. Power output follows the quadratic V^2/R relationship. Efficiency performance falls off drastically as the amplifier exits the Class-E operating mode. Retuning the load for these voltages is possible but would require considerable effort to recalculate the load and reoptimize for each level.

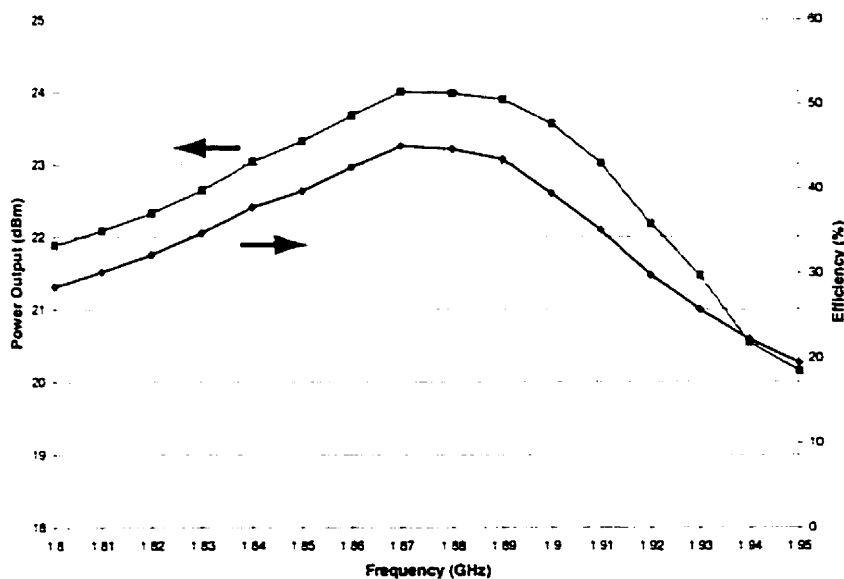


Figure 2.26: Predicted Performance Versus Frequency

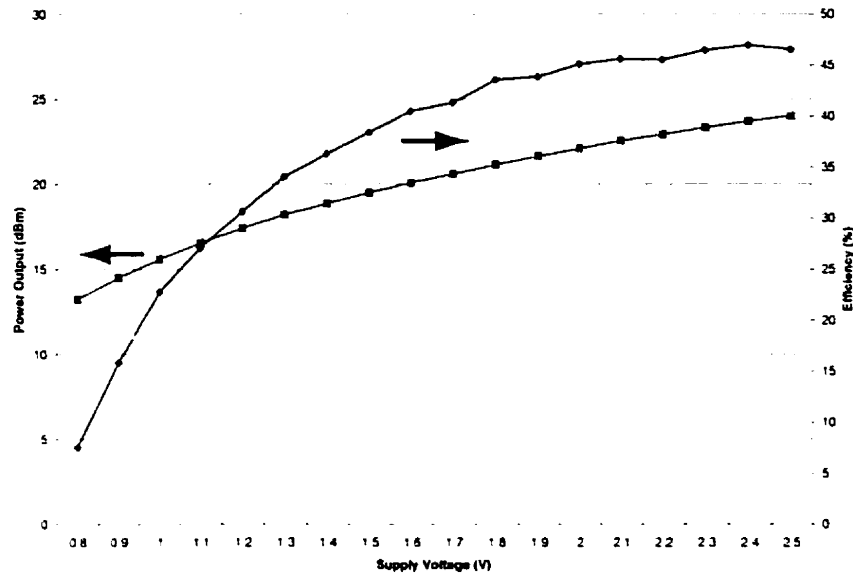


Figure 2.27: Predicted Performance Versus Supply Voltage

2.13 Summary

The theoretical basis behind the design of the Class-E amplifier has been discussed. A detailed discussion of the considerations involved with modeling the output stage was presented in the first section of this chapter. The complete narrowband design detailed some of the work necessary to achieve reasonable performance at 1.9 GHz.

The design of an appropriate gain stage capable of driving the output Class-E amplifier was also presented. The rationale and design considerations for this stage were also discussed.

Design considerations involving input matching were presented in this chapter. namely the design of the on-chip spiral inductor for use in the matching circuit were rigorously investigated.

The chapter concluded with a presentation of the overall circuit and with simulated results for the complete design. Peak output power and efficiency were both predicted to be 24 dBm and 45% respectively.

2.14 References

- 1 N. Sokal, A. Sokal, "Class E – A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 10, pp. 168-176, 1975.
- 2 F. H. Raab, "Idealized operation of the Class E Tuned Power Amplifier," *IEEE Transactions on Circuits and Systems*, vol. 24, pp. 725-735, 1977.
- 3 R. E. Zulinski, "The effects of parameter variations on the performance of diode-modified Class-E tuned power amplifiers," M.A.Sc. thesis, Michigan Technological University, 1980.
- 4 R. E. Zulinski, "Class-E frequency multipliers," Ph.D. dissertation, University of Wyoming, 1985.
- 5 R. E. Zulinski, J. W. Steadman "Class E power amplifiers and frequency multipliers with finite DC-feed inductance," *IEEE Transactions on Circuits and Systems*, vol. 34, pp. 1074-1087, 1987.
- 6 R. E. Zulinski, K. J. Grady "Load-independent class E power inverters: part I – theoretical development," *IEEE Transactions on Circuits and Systems*, vol. 37, pp. 1010-1018, 1990.
- 7 C. H. Li, Y. O. Yam, "Maximum frequency and optimum performance of class-E power amplifiers," *IEE Proceedings on Circuits Devices and Systems*, vol. 141, pp. 174 - 184, 1994.
- 8 J. J. Ou, X. Jin, I. Ma, P. R. Gray, "CMOS RF modeling for GHz communication ICs," *Symposium on VLSI Technology*, pp. 94-95, 1998.
- 9 R. A. Pucel, "Design considerations for monolithic microwave circuits," *IEEE Transactions on Microwave Theory and Techniques*, vol. 29, pp. 513-534, 1981.

-
- 10 G. Eleftheriades, Microwave Techniques, University of Toronto, Lecture Notes, 1998.
 - 11 K. Tsai, P. R. Gray, "A 1.9 GHz, 1-W Class E power amplifier for wireless communications." *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 962-970, 1999.
 - 12 T. Sowlati, Y. Greshishchev, and C.A.T. Salama "1.8 GHz Class E Power Amplifier for Wireless Communications." *Electronic Letters*, vol. 32, pp. 1846-1848, 1996.
 - 13 J. Doyama, "CMOS Class E Power Amplifier." M.A.Sc. Thesis, University of Toronto, 1999.
 - 14 D. A. Johns, K. Martin, Analog Integrated Circuit Design, John Wiley & Sons, 1997.
 - 15 F. Herzel, B. Razavi, " Oscillator jitter due to supply and substrate noise." *Proceedings of the IEEE CICC*, pp. 129-132, 1995.
 - 16 N. M. Nguye, R.G. Meyer, "Si IC-compatible inductors and LC passive filters." *IEEE Journal of Solid-State Circuits*, vol. 25, pp. 1028-1031, 1990.
 - 17 K. Negus, B. Koupal, J. Wholey, K. Carter, D. Millicker, C. Snapp, N. Marion, "Highly-integrated transmitter RFIC with monolithic narrowband tuning for digital cellular handsets." *ISSCC Digest of Technical Papers*, pp. 38-39, 1994.
 - 18 J. Y. C. Chang, A. A. Abidi, M. Gaitan, "Large suspended inductors on silicon and their use in a 2- μ m CMOS RF amplifier." *IEEE Electron Devices Letters*, vol. 14, pp. 246-248, 1993.
 - 19 A. Sedra, K. Smith, Microelectronic Circuits, Saunders College Publishing, 1991.
 - 20 G. D. Vendelin, A. M. Pavio, U. L. Rhode, Microwave Circuit Design, John Wiley and Sons, 1990.

-
- 21 Guillermo Gonzalez. Microwave Transistor Amplifiers. Prentice Hall, 1984.
 - 22 H. A. Wheeler. "Simple Inductance Formulas for Radio Coils." IRE Proceedings. p. 1398, 1928.
 - 23 J. P. Holman. Heat Transfer. Prentice Hall. 1986.
 - 24 W. S. Janna. Engineering Heat Transfer. PWS Publishers. 1986.
 - 25 T. S. Hashjani. "Class-E power amplifiers for wireless communications." Ph.D. Dissertation, University of Toronto. 1996.

CHAPTER 3

Class-E Power Amplifier Layout

3.1 Introduction

Power amplifiers are amongst the most power-hungry building blocks in an RF transceiver. The tremendous current levels and high slew rates are the main reasons behind the difficulties associated with designing, and especially packaging these components. Parasitic resistances on the order of tens of milliohms and inductances on the order of tens of picohenries may result in considerable efficiency losses. For these reasons, many layout and packaging issues that are usually unimportant in other analog and RF circuits become crucial in power amplifiers [1].

3.2 Transistor Layout

A unit cell for the output device was initially designed. Each cell was designed to be the equivalent of a transistor that is 500 μm in channel width. Each of these cells actually consists of 20 individual transistors that are 25 μm in width. This type of layout is often referred to as an interdigitated layout. A plot of one of these interdigitated transistors, as implemented, can be seen in Figure 3.1. The main purpose of an interdigitated layout is a reduction in area that a large channel width transistor would otherwise occupy. For further information, there are numerous references that

would serve as a design guide. The text by Johns and Martin is a good reference in this regard because it concisely summarizes a number of these techniques [2].

Dividing up large transistors also improves transient time. Since modern MOSFET gates are constructed from highly resistive polysilicon and since gate signals can only be injected at either ends of the device, it is apparent that long channel width devices will experience slewing difficulties as a result of RC time constants. This is especially prevalent when operating frequencies approach the transition frequency, f_t . It is important to be aware of this problem because current BSIM 3 and Level 28 transistor models do not account for either the gate resistance or drain to bulk peripheral resistance as mentioned previously. The effects of these elements become more significant at frequencies above 2 GHz. When these factors are considered, the application of an interdigitated layout becomes necessary to minimize channel width.

As discussed in the previous chapter, channel length was selected to be the minimum possible. A minimum channel length is desired for the fastest possible device switching speed, f_t , and ultimately the optimal overall performance. The minimum available channel length of 0.35 μm was therefore chosen for both the gain and output devices at the time of implementation.

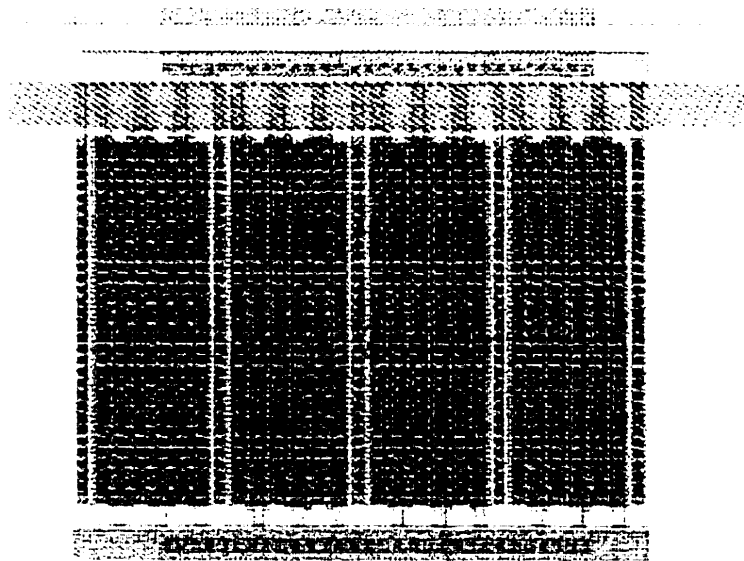


Figure 3.1: Transistor Unit Cell Layout

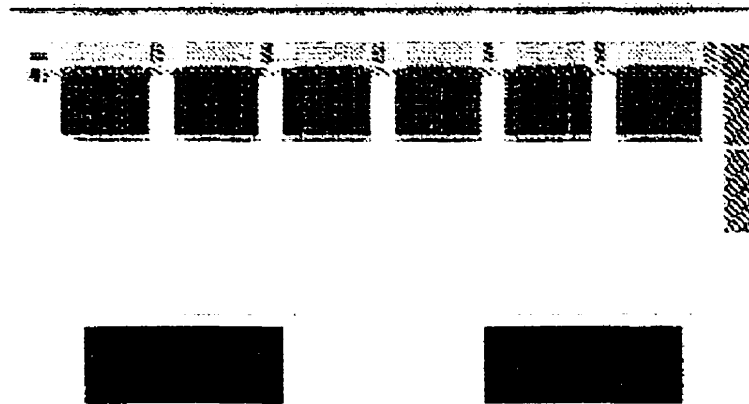


Figure 3.2: Unit Cell Array for Output Device

A combination of six of these unit cells was used to create the 3000 μm output device. The diagram of these cells as configured for use in the output stage is shown in Figure 3.2. Since most of the power dissipation in an RF amplifier occurs in the actual channel of the device, spreading these cells over a larger area benefits heat dissipation. The unit cells in this case were spread as far apart as possible without impacting the overall layout.

There are no detriments to placing many contacts to connect the $n+$ drain and source diffusion regions to the Metal 1 layer. Placing a large number of contacts will only serve to benefit the result by minimizing the drain and source resistance associated with this interface.

It was also not desirable to place the bondpads too close to the devices. Since the bondpads and metal interconnects were only implemented in the top metal layers, the stresses encountered during the wirebonding process could have easily damaged the fine-geometry devices if the bondpads were placed too close. Excessively long interconnects should also be avoided though to prevent unnecessary attenuation or ground coupling of large RF signals.

3.3 Capacitor Layout

Figure 3.3 is a layout of the interstage coupling capacitor used. The interstage coupling capacitor was made large to minimize the series impedance between the two devices. The leads connecting the device to the transistors were made as short as possible in order to minimize any possibility of signal attenuation. Numerous contacts were also used in this case to aid capacitor coupling to the interconnects.

A polysilicon-polysilicon capacitor was chosen in lieu of a metal to metal capacitor because although the metal-to-metal capacitors have very good resistive properties, the poly-poly caps have much higher capacitance per unit area. A smaller area would therefore result in less area consumed and a reduced coupling to ground.

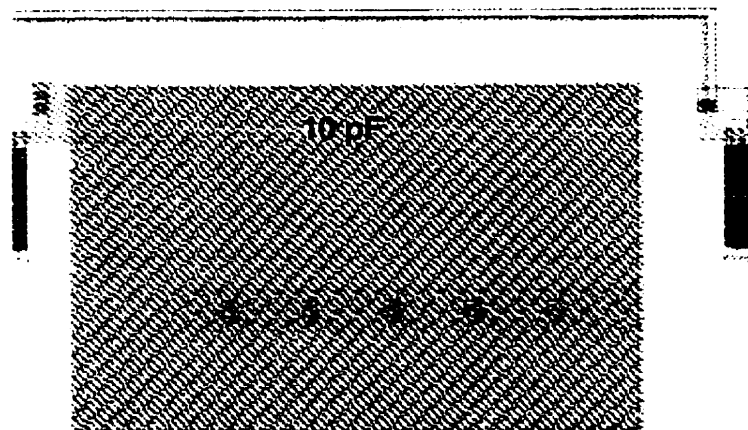


Figure 3.3: Interstage Coupling Capacitor

3.4 Grounding Vias

Grounding vias were spaced away from the transistors to prevent excessive coupling to the ground plane [3]. This effect is illustrated in Figure 3.4. When grounding vias are placed near the drain region of the transistor, this reduces the

amount of impedance between the drain periphery and ground. This would eventually increase the amount of parasitic loss seen at the drain at high frequencies.

In this implementation, substrate contacts were generously placed throughout the ground plane field. This was to ensure good coupling of the substrate to ground. Although this ultimately served to reduce substrate current levels, in a fully integrated transceiver design a p+ guard ring would also be implemented around the perimeter of the power amplifier circuit. Although this project was specifically implemented in a differential form to minimize substrate currents, some residual leakage will always occur. This has also been shown to be especially true at frequencies above 2 GHz. All possible measures must therefore be taken to safeguard receiver circuitry as indicated by Hansen [4]

Finally, it is important to be aware that many processing steps have been devised to address grounding in high-speed circuit design. Wrap-around grounding and plated-hole vias have been successfully used to aid both grounding and heat conduction to the ground plane on the back the die [5][6]. It is anticipated that these techniques will soon be necessary in volume production processes to increase circuit performance. If these features do become available, they should be used in RF amplifier design.

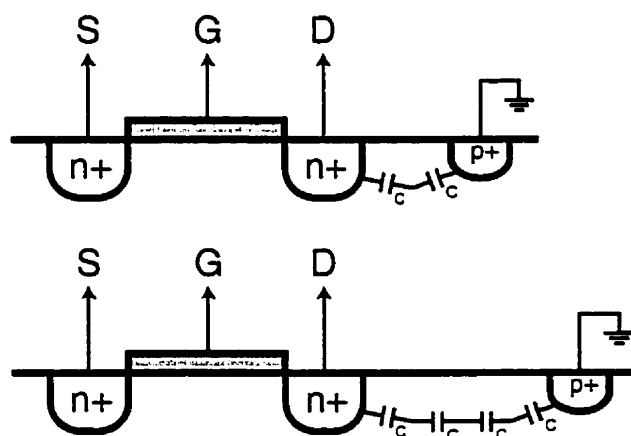


Figure 3.4: Grounding Via Placement

3.5 Inductor Layout

There have been a number of highly exotic processes developed specifically to optimize the performance of RF inductors [7]. The general idea behind these techniques is to reduce the capacitance of an inductor's windings to ground. Low resistance gold metallization also arrives at the same result. Although none of these additional processing steps were available in the conventional CMOS technology used for implementation, steps can be taken to optimize performance.

Figure 3.5 is an illustration of the layout used for the input matching inductor. This inductor was implemented using the top metal layer only. Although the series resistance for the spiral will be higher than if multiple metal layers are used, the metal windings are farther away from the conductive substrate and will therefore experience less shunt capacitance. Also, since accurate modeling of these inductors is a key issue, the higher predictability of one metal routing layer was seen as a significant benefit.

The windings of the inductor are not continued fully to the center of the coil. This allows fields to circulate through the center and around the windings. Allowing the fields to circulate prevents current crowding in the inner windings of the inductor. This is caused by fields which are forced through an opening that is too small. This effect results in a higher series resistance seen in these windings. In the end, overall inductor quality factor is reduced. This phenomenon was proven by Lin et al [8].

As well as current crowding, other layout issues must be considered when designing an on-chip spiral inductor. Long and Copeland documented a number of details to consider [9]. These additional parameters are summarized in the following:

- Maintain at least 5 linewidths of space between the outer turns of the spiral and any surrounding metal features.
- Spacing between adjacent metal lines of the spiral should be minimized to optimize line to line magnetic coupling.
- Strip widths should be chosen between 10 to 15 μm for present day, state-of-the-art semiconductor processes.
- The oxide layer which isolates the metal conductors from the silicon substrate should be kept as thick as possible to minimize shunt parasitics and power dissipation.

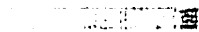


Figure 3.5: Input Matching Spiral Inductor

All of the mentioned techniques mentioned by Long were applied to optimize the performance of the device.

3.6 Bondpad Layout

Special considerations were made when implementing the bondpads used in the RF signal path. The typical bondpads used for conventional digital circuits, or for low speed analog implementations are actually complex devices. These devices are intended to shield the IC core from such effects as electrostatic discharge. Unfortunately, at 2 GHz these devices are highly parasitic. In RF designs where efficiency performance is key, these devices should therefore be avoided.

In addition to removing the bondpad protection circuitry, the lower layers of the bondpads were also stripped away to minimize coupling to the substrate. Although this makes the more bondpads more delicate during the bonding process, this step is necessary to optimize performance. For added substrate isolation, an n+ well was implanted under the bondpad. This is similar to the ground shielding technique implemented by Meyer [10].

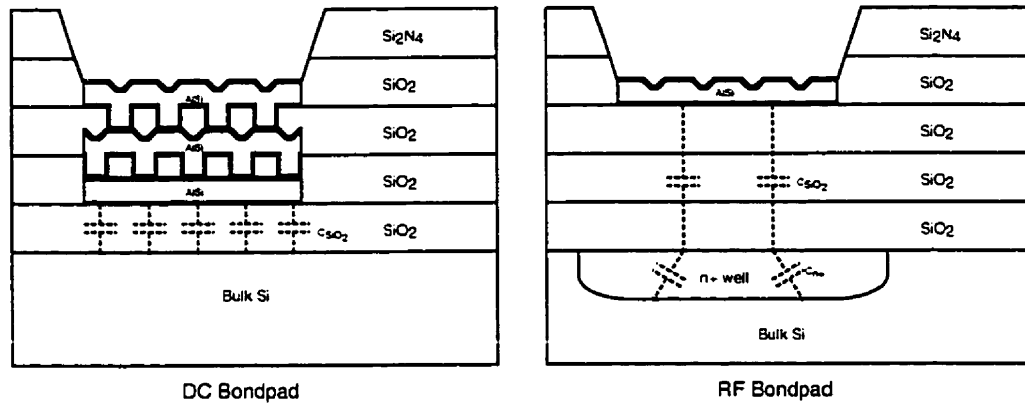


Figure 3.6: Comparison of DC and RF Bondpads

Figure 3.6 is an indication of how these modifications will come into play when compared to a conventional bondpad. As seen in the illustration, the RF bondpad is farther away from the substrate and is therefore less capacitive. In addition, the n+ implantation serves as another capacitor placed in series with the oxide capacitance. These techniques combine to reduce the overall capacitance of the bondpad to the substrate.

3.7 RF Signal Paths

RF signal paths are exclusively run in the top metal layer. This serves to minimize the parasitic capacitance to the substrate. Distributed transmission line analysis was not necessary at the desired frequency because of the small chip dimensions involved. One-tenth of a wavelength for an interconnect should serve as a point where distributed analysis is necessary.

3.8 Ground Plane Considerations

The ground plane was made large and spans all three available metal layers. This was primarily to satisfy metal fill requirements for the process. Good coupling between the ground and substrate is necessary to prevent oscillation. Substrate contacts were placed throughout the field to minimize any potential for oscillation.

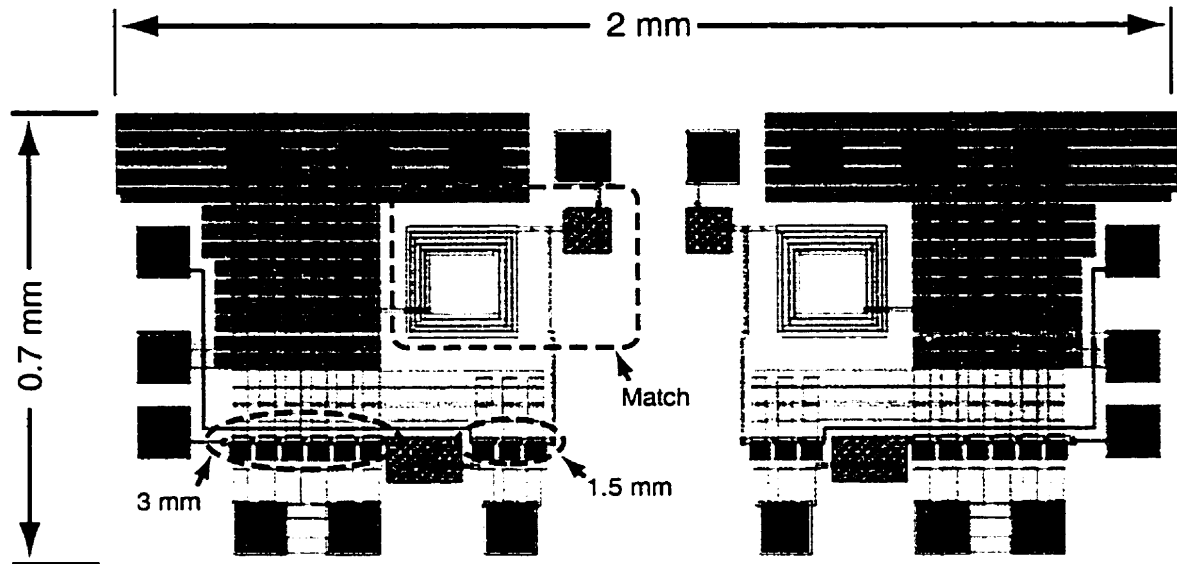


Figure 3.7: Final Amplifier Layout

3.9 Final Implementation

Figure 3.7 is a plot of the final amplifier as submitted for fabrication. The active chip area in this case is indicated to be 1.4 mm^2 . A majority of this area is occupied by a passive ground plane. Test package inductance and the number of leads required for grounding largely dictated the number of bondpads required. This ultimately determined the general layout of the design presented here.

The differential circuit is designed to be a mirror image from left to right. The symmetry maximizes the noise cancellation benefits when driven differentially. Implementing non-interconnected amplifiers aided testing and development. Since each half of the amplifier could be tested independently this allowed for easier diagnosis of any problems encountered.

3.10 Conclusions

This chapter summarized some of the layout considerations necessary for the amplifier's design. The amplifier was implemented in TSMC's triple-metal, double-poly, $0.35 \text{ }\mu\text{m}$, process. As discussed, the layout of RF designs and especially high

power amplifiers require consideration well beyond what would be expected for low power circuits.

3.11 References

- 1 B. Razavi, "Recent advances in RF integrated circuits." *IEEE Communications Magazine*, December, pp. 36-43, 1997.
- 2 D. Johns, K. Martin. *Analog Integrated Circuit Design*, John Wiley and Sons, 1997.
- 3 Z. Zhang, A. Pun, J. Lau, "Interference issues in silicon RFIC design." *IEEE Radio Frequency Integrated Circuits Design Symposium*, vol. 1, pp. 119-122, 1998.
- 4 K. Hansen, "Wireless communications devices and technology." *IEEE Radio Frequency Integrated Circuits Symposium*, vol. 1, pp. 1-5, 1998.
- 5 R. A. Pucel, "Design Considerations for Monolithic Microwave Circuits." *IEEE Transactions on Microwave Theory and Techniques*, vol. 29, pp. 13-34, 1981.
- 6 T. Ruttan, "Designing amplifiers for wireless systems." *Microwaves and RF*, January, pp. 89-100, 1998.
- 7 J. S. Kim, C. H. Park, S. H. Kim, G. H. Ryu, K. Seo, "F-Inductor and BC-MOS technology for monolithic silicon RF ICs." *IEEE Radio Frequency Integrated Circuits Symposium*, vol. 5, pp. 165-168, 1998.
- 8 H. Tsai, J. Lin, R. C. Frye, K. L. Tai, M. Y. Lau, D. Kossives, F. Hrycenko, Y. Chen, "Investigation of current crowding effects in spiral inductors." *IEEE Microwave Technologies and Techniques Symp.*, pp. 139-142, 1997.
- 9 J. Long, M. Copeland, "The modeling, characterization, and design of monolithic inductors for silicon RF ICs." *IEEE Journal of Solid State Circuits*, vol. 32, pp. 357-369, 1997.
- 10 R. G. Meyer, W. D. Mack, "A 1-GHz BiCMOS RF front-end IC." *IEEE Journal of Solid State Circuits*, vol. 29, pp. 350-355, 1994.

CHAPTER 4

Experimental Results

4.1 *Experimental Implementation*

A photomicrograph of the implemented amplifier can be seen in Figure 4.1. Before testing electrically, the devices were inspected visually. Occasionally, dicing dies from a wafer will produce fractures in the die. Sometimes these fractures are not clearly visible, even when viewed with a microscope. Varying light and magnification levels were therefore used to thoroughly scrutinize the devices before performing any tests.

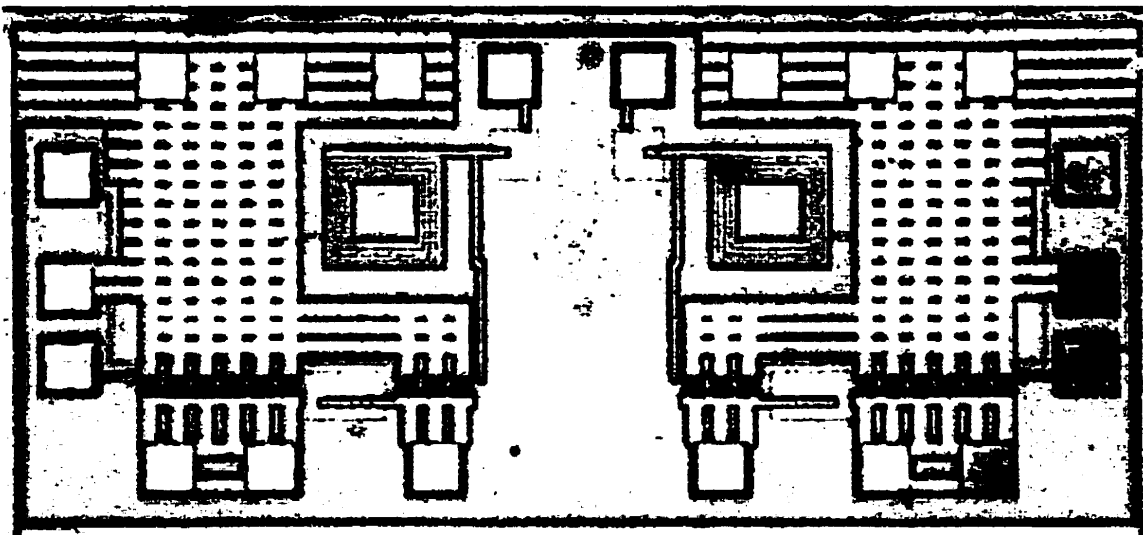


Figure 4.1: 1.9 GHz Amplifier Photomicrograph

4.2 DC Testing

Initially, transistor functional parameters were verified with multimeter measurements. A Fluke 87 multimeter was used throughout this process. Being battery operated and highly accurate, the output sense voltage produced by this unit was confirmed to be at a benign level and within device operating allowances. This is a necessary precaution because some lab bench multimeters will produce sense voltages that can easily damage fine-line geometry transistors. This precaution is especially relevant because of the lack of any bondpad protection circuitry.

Additional test points were included in the design to allow the transistor gates to be probed. A MOSFET gate will typically fail by shorting to ground. A high impedance gate measurement was therefore seen as an indication that the thin gate oxide was in good condition. A number of devices were actually confirmed to be faulty using this technique.

Saturation current was also measured during DC testing. The saturation current value served as a final verification that the devices were operating within proper tolerances. Occasionally, a device was seen to conduct current even when a finite gate resistance indicated a damaged transistor. In all of these cases though, when saturation current values deviated significantly from what was predicted, these chips were deemed to be faulty.

4.3 Low Power RF Testing

In comparison to the solder-attached tungsten-copper heat spreaders used in high power RF implementations, the best available alumina package with its conductive-epoxy interface was a poor conductor of thermal energy in comparison [1]. The half Watt of power required for dissipation in this project under typical conditions would have resulted in device failure. In addition to the characteristically high thermal resistance, the inability to directly attach a heat sink to this package was another detriment to the configuration. When testing at higher power levels was actually attempted with this package, device failure did, in fact, occur.

High power testing is also inappropriate with this package because of the lossy package lead-frame. In comparison to other, more optimal techniques, efficiency measures would have suffered. Package parasitics also mean a larger ground-plane impedance. In addition to the reduced amplifier drive signal discussed previously, the potential oscillation would have also becomes more likely as power levels are increased.

4.4 Low Power Test Results

Rapidly sweeping the amplifier's input signal over a broad band is a valid test of stability. This is actually similar to the way in which a network analyzer probes components. A network analyzer was therefore used to first test the device at frequency. Sweeping the input signal from 1.6 to 2 GHz and observing all four of the two-port s-parameters confirmed that the device was functioning properly. Any large spikes in either of the transmission or reflection magnitudes would have been a positive indication of any oscillatory behavior.

Evaluation with a spectrum analyzer is essential. It is virtually impossible to use an oscilloscope to troubleshoot RF circuits. Any off-order harmonics will generate an output signal that is virtually unrecognizable on an oscilloscope display and actually prevent the scope from triggering. This was the case when the device was initially tested. When the signal was examined with an HP 8653 spectrum analyzer though, the problem became clear. It was readily apparent that some spurious signals below the fundamental were causing the problem. Although the magnitude of these signals was low, the addition of some supply decoupling reduced these signals to a more acceptable level.

Figure 4.2 through Figure 4.4 shows the spectrums produced during the frequency sweeps. As seen in these plots, the outputs were noted to be clean and free from any significant spurious signals. Low harmonics levels were used as an indication that the amplifier was operating in a linear manner.

The power levels involved here are too small to fully invert the large channel-width devices. As a result, gain and output power were small. The output stage was also unmatched during these tests. Since this amplifier was not intended for broadband applications, optimal power transfer was therefore impossible. Attempting to drive the amplifier at high power levels would have led to standing waves that could have caused

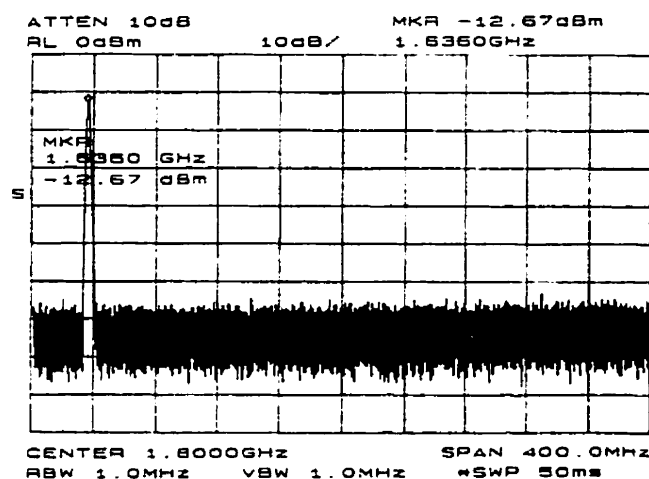


Figure 4.2: 1.6 GHz Low Power Spectrum

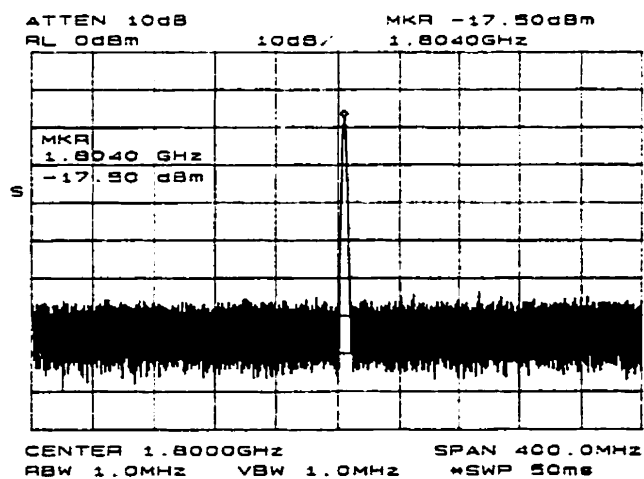


Figure 4.3: 1.8 GHz Low Power Spectrum

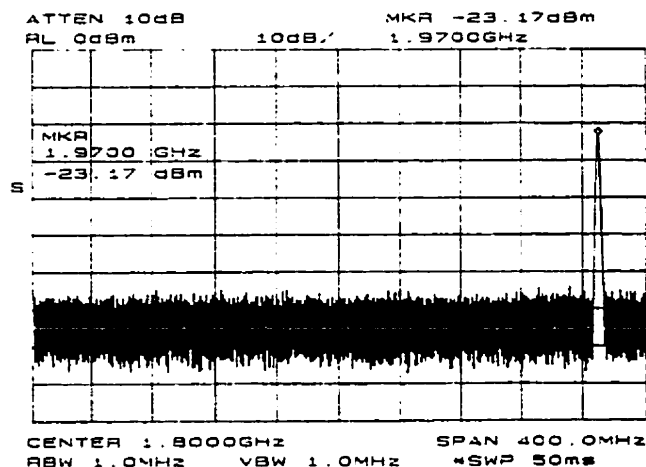


Figure 4.4: 2 GHz Low Power Spectrum

device failure. Frequency sweeps were therefore strictly used as a functionality measure.

4.5 High Power RF Testing

It quickly became apparent that the best package available at the time of the implementation was still poor by high-power RF standards. Although the packages were advertised as having a 2 GHz operating frequency, heat-transfer calculations proved that this component would have been problematic. The heat generated in this application therefore required special consideration. This conclusion was also made by a number of other researchers [2.4]. The high resistance levels in the input and output signal paths would also have affected overall performance.

Another problem with this package-circuit board combination was that there were no calibration fixtures available. Calibration fixtures are necessary when doing any type of S-parameter analysis. Without a calibration structure, the designer is unable to account for the impedances and phase delays associated with a given test setup. Because of this inability to calibrate, there was no way of validating the package models associated with the design. This also prevented the accurate determination of the true input match. Additionally, even if the supplied boards could have been characterized there was still the interaction of the board to the package that would not have been

considered. Positional differences would have induced phase delay errors, invalidating the results.

There are other issues with the supplied circuit boards that would have led to reduced performance. The poor ground-plane design would have resulted in higher ground impedance levels than originally anticipated. Since the exact characteristics of the board design were unknown, its effects on the circuit would have been immeasurable.

In addition to the above limitations, a manufacturing defect meant that a number of the leads used for grounding were virtually useless. Given the extremely sensitive nature of the design to layout and packaging, this problem meant that the boards were unacceptable.

To resolve the mentioned packaging issues, a number of other designers have adopted a technique whereby the bare die is mounted directly onto the printed circuit board substrate [5]. This approach is appropriately referred to as chip-on-board mounting.

By specifically tailoring a printed circuit board layout to a given application, there are a number of benefits that can be gained. In general, all of the parasitic losses associated with packages can be entirely avoided. Input and output 50-ohm transmission lines can be directly connected to the signal-path bondwires. The length of these bondwires can also be reduced from what would otherwise be possible. As well, decoupling capacitors can be located closer to the chip to help minimize any potential for oscillation.

FR4 material was used for a custom chip-on-board design. FR4 is widely available but is lossy above 2 GHz. This material was available in 30 mil thicknesses and relative permittivities of 3.5. Calibrated measurements confirmed these parameters.

A number of factors combined to contribute to a key lower ground-plane impedance. Shorter grounding bondwires were used to attach directly to wide low-impedance ground traces. Strategically located vias were implemented to connect the top side ground traces to the back side ground plane. The board material was also chosen to be the thinnest possible to minimize the length of these vias. An added benefit of the thin boards were narrower 50-ohm transmission line widths. This helped to reduce the

amount of taper necessary to access the 2 mm wide die. Finally given the precise control over the design parameters, the performance of all of these elements could be predicted by using electromagnetic solvers.

Heat transfer was also greatly improved with this configuration. Vias were specifically located under the die to act as heat pipes. As described in [6], heat pipes can be used to significantly reduce the thermal resistance of a circuit board. Thermally and electrically conductive silver epoxy was used to form a bond with the top side circuit board ground plane. Finally, a virtually infinite heat sink in conjunction with heat sink compound were used as the final step in minimizing thermal resistance.

Figure 4.5 is a schematic representation of the chip-on-board test fixture. The gray area indicates the copper used for the top metal layer. The pairs of wide input and output traces represent the widths necessary to achieve 50-ohm transmission lines. The narrow circuit-board traces are quarter-wave meander lines used for biasing.

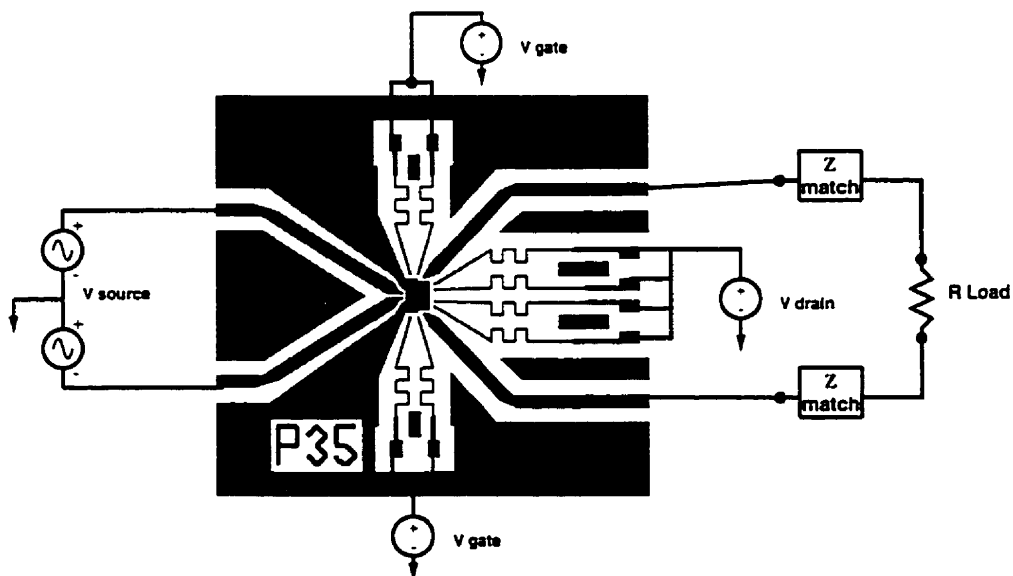


Figure 4.5: Schematic Representation of Chip-On-Board Test Fixture

A photograph of the test setup is shown in Figure 4.6. The photograph clearly indicates where the launches are positioned to interface with standard SMA cables. The dark plane beneath the fixture is the black-anodized surface of the large heatsink.

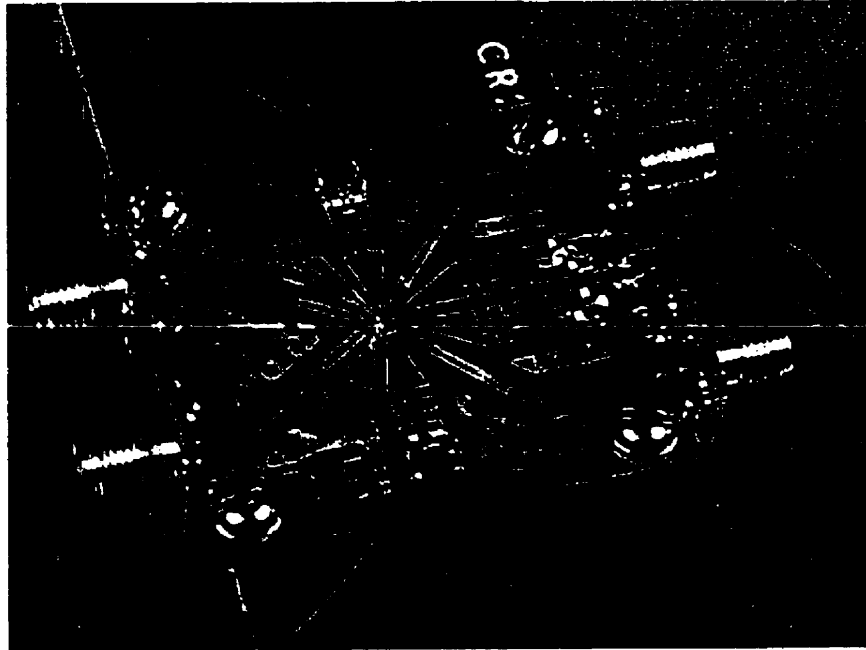


Figure 4.6: Implemented Design Showing Heatsink and Launches

To facilitate the accurate extraction of device parameters it is absolutely essential to have a calibration structure. Without such an apparatus, it is virtually impossible to characterize performance parameters such as input return loss or power gain. These calibration structures are typically used when making any type of measurement involving a network analyzer. These devices can also be used in high power test setups to determine setup losses.

The calibration structure developed for the test fixture is shown in Figure 4.7. This structure is intended for a SLOT network analyzer calibration. The acronym SLOT stands for short, load, open, and through. Short open and load refers to the impedance presented to the ends of the transmission lines identical to the ones actually used for testing. In this way, all of the parasitics associated with the cables, connectors and the test fixture itself can be extracted from the final measurements. Upon closer examination of the diagram, it is apparent that each of the calibration standards is, in fact, one half of the symmetric circuit board.

The short and open standards are used to give an accurate zero reference in the calculation of proper phasing in reflection measurements. The load standard is used in

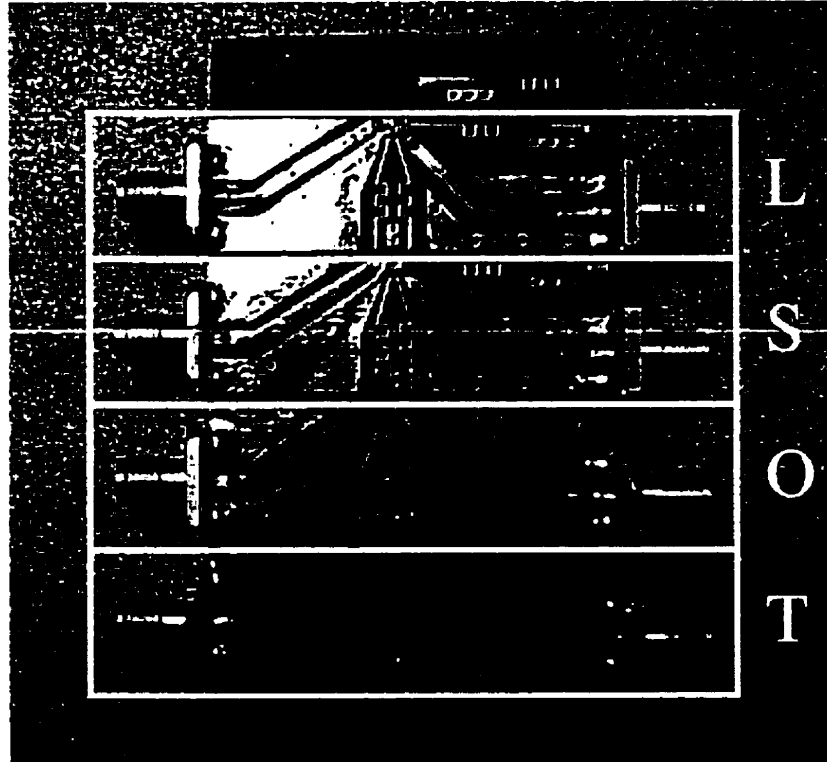


Figure 4.7: Chip-On-Board Calibration Structure

conjunction with the short and open standards to determine accurate return loss magnitudes. The through standard is used to determine both the forward and reverse transmission losses as well as the respective phase delays.

4.6 High Power Test Results

As with the low power testing, a HP 8653 spectrum analyzer was used to diagnose the chip-on-board mounted amplifier. With this approach, it was determined that the system was resonant at 1.3 GHz. Through trial and error, it was determined that the bias lines were the cause of the resonance. The addition of shunt capacitance on these lines, though, remedied this problem.

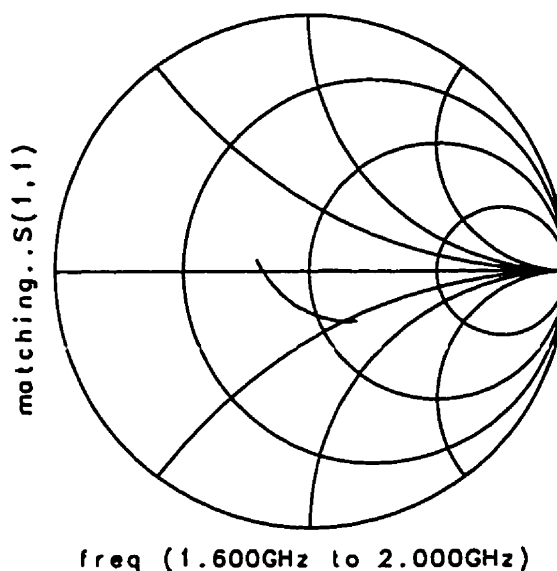


Figure 4.8: Input Reflection Coefficient

The amplifier's input match was initially measured with the available HP 8720B vector network analyzer. The calibration structure was therefore used to perform a SOLT calibration before testing began. The results of these measurements are shown in Figure 4.8 and Figure 4.9. Figure 4.8 is a Smith Chart plot of the input reflection coefficient, S_{11} . This plot indicates a smooth and tractable phase progression of the input match. Figure 4.9 displays the magnitude of this input match.

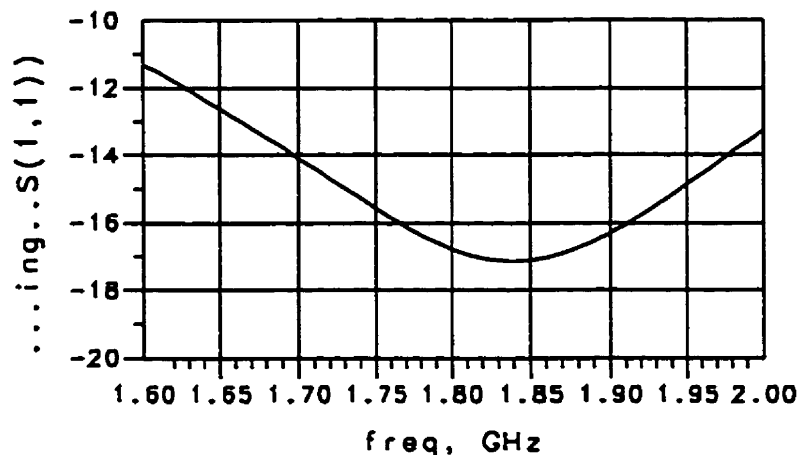


Figure 4.9: Input Reflection Magnitude

As seen in Figure 4.9, although the center frequency of the input match was approximately 5 MHz below the desired 1.88 GHz transmit-band center frequency, a better than 12 dB match was achieved across the entire 1.85 to 1.91 GHz transmit bandwidth. This served as an indication that the input matching network was functioning properly.

Much work was done to develop a high power testing methodology. The culmination of this experimental work is the test setup shown in Figure 4.10. In lieu of a prohibitively expensive load-pull measurement machine, this setup was used as a low cost alternative. Load pull testing was primarily used to determine the load impedance for optimal full-power performance.

The network analyzer was used as both a signal source and a probe for the output reflection measurements in the load-pull setup. The variable gain amplifier provided 20

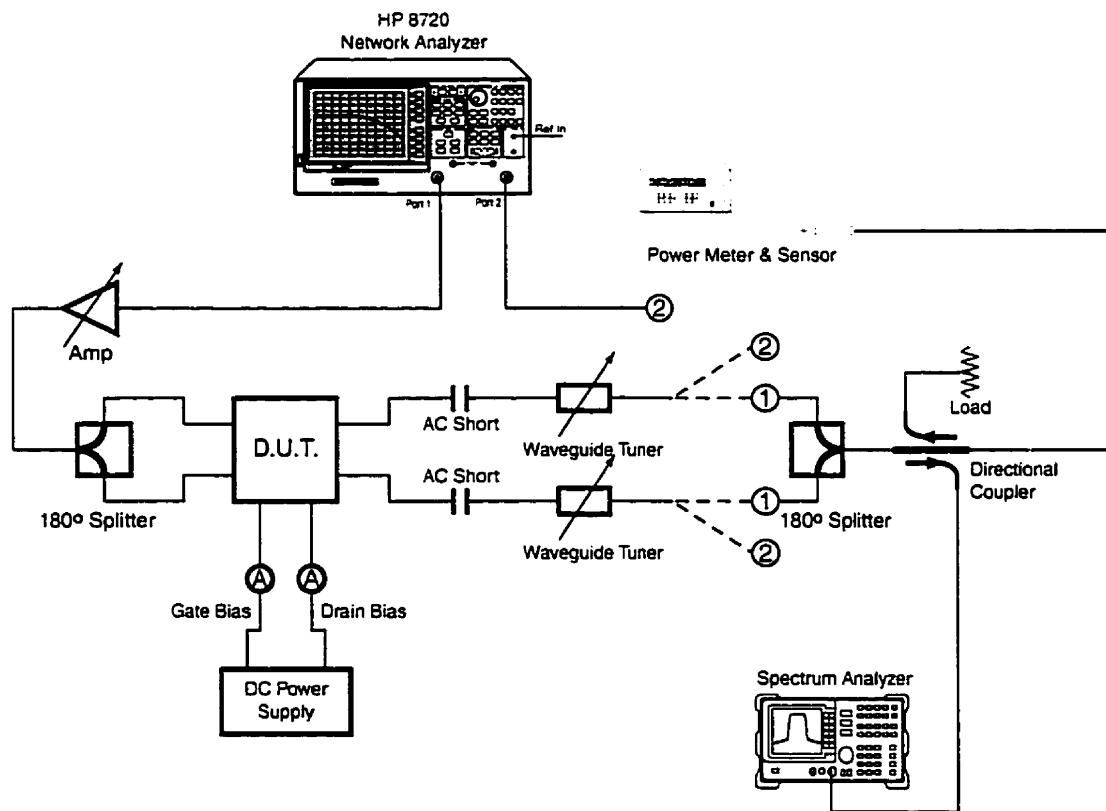


Figure 4.10: Load Pull Analysis Test Setup

dB of gain for the 8720's -10 dBm output signal. 180° microstrip phase splitters, having no more than 4 degrees of phase error, were used for the single ended to differential conversion. DC blocking capacitors were used to prevent the waveguide tuners from shorting-out the output devices.

The principle behind the load pull tuning is straightforward. Waveguide tuners are used to change the impedance presented to the output device. The calibrated network analyzer is then used to measure the impedance presented to the device. Once an impedance level is set, leads are disconnected from the network analyzer (2) and then connect the power splitter (1). A power meter is then used to determine the output power generated for that particular impedance level. This process is iterated for impedances covering the entire Smith Chart. As a result of this procedure, contours of constant power, for instance, can be generated on the impedance plane of the Smith Chart. Results will resemble those produced by Maeng et. al [7].

Enhancements to the load pull setup entailed the addition of a spectrum analyzer. The integration of a spectrum analyzer in the test setup permitted the observation of the output spectrum while the amplifier was being tested at full power. In this way, a number of resonances that only appeared at peak power were discovered and eliminated. If these parasitic oscillations were left undiscovered, maximum output power would have been restricted. Using a directional coupler to probe the output signal meant a minimal impact on load impedance and power measurement accuracy. A power meter should be used for these types of measurements in lieu of the network analyzer because the power levels involved would damage a typical analyzer.

Figure 4.11 through Figure 4.13 show the final output results from the amplifier measurements. In general, results were similar to those predicted. A problem with chip dicing resulted in additional material being left on the end of the die. This caused the critical bondwires used for drain loading to be longer than anticipated. The end result of this was a reduced resonant frequency and an ultimate efficiency loss. The effects of these non-ideal bondwires can be seen in the experimental results shown in Figures 4.11 through 4.14. This cause of this efficiency loss was verified through simulation. This should also serve as an indication of the sensitivity of the design to implementation.

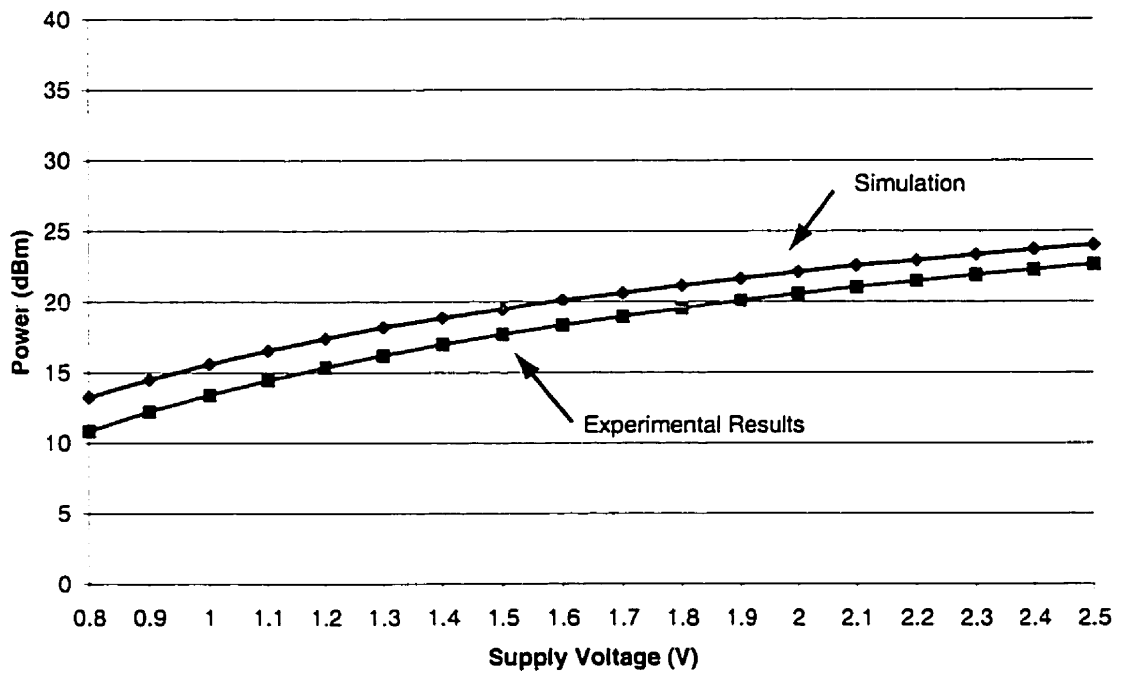


Figure 4.11: Output Power vs. Supply Voltage

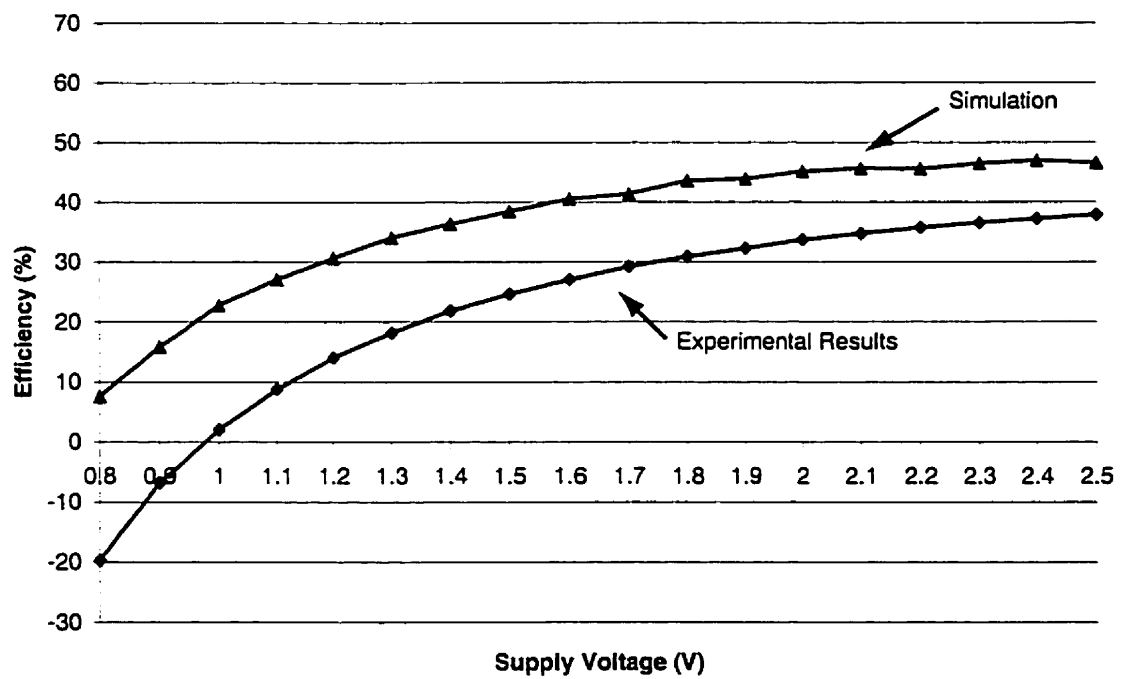


Figure 4.12: Efficiency vs. Supply Voltage

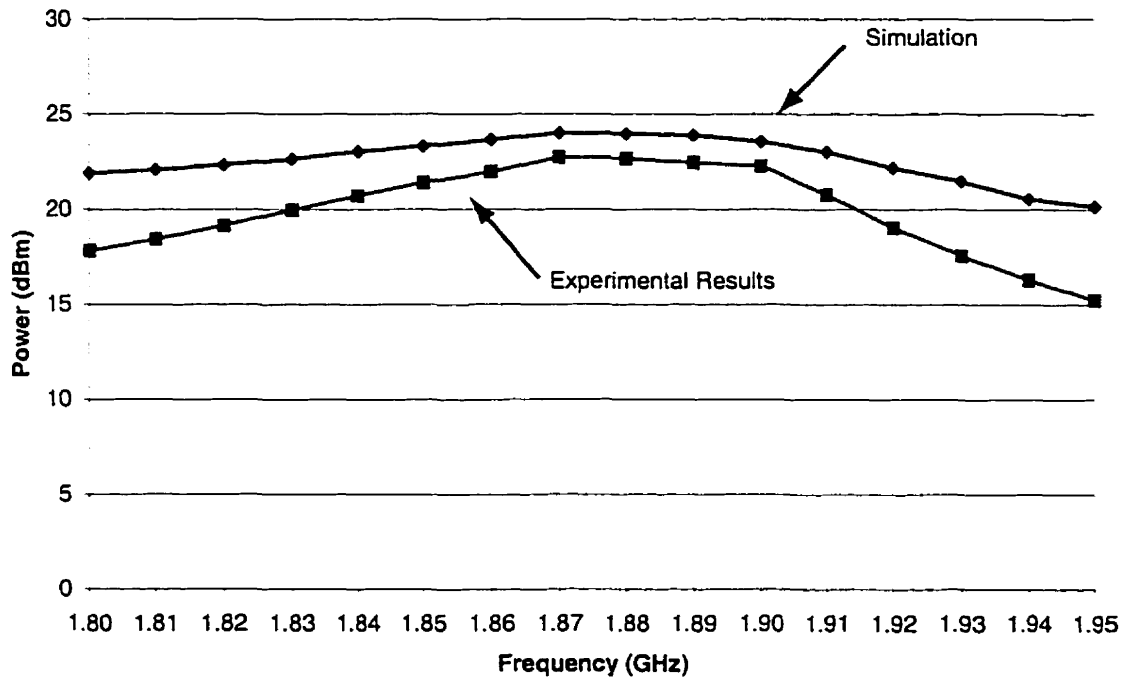


Figure 4.14: Output Power vs. Frequency

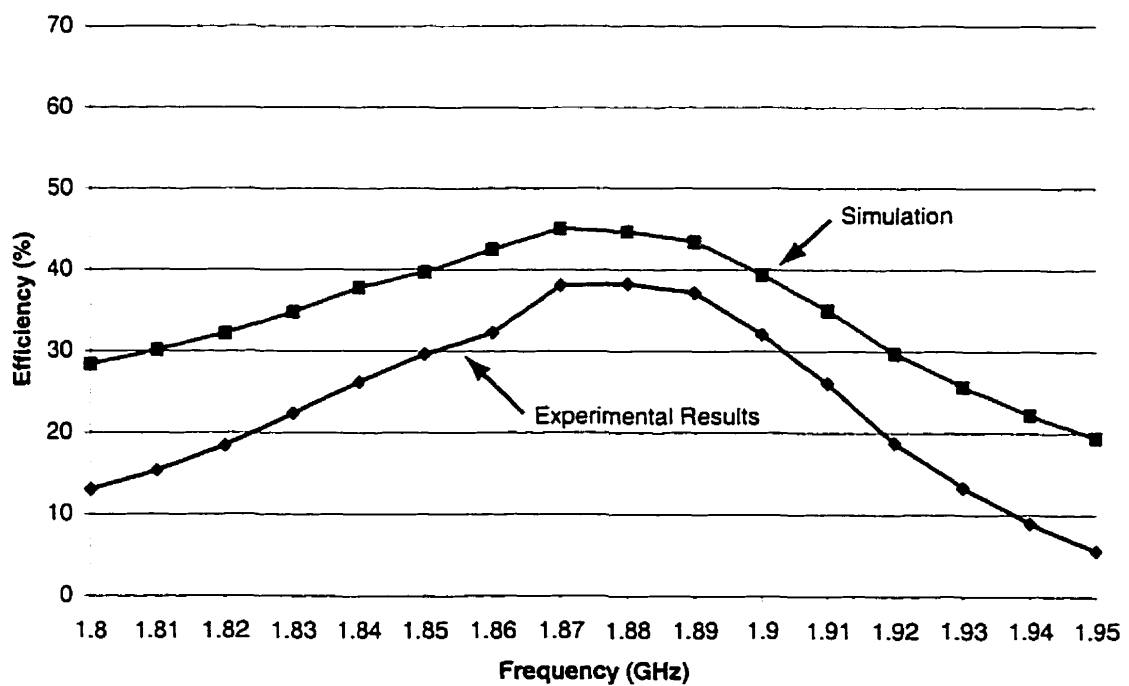


Figure 4.13: Efficiency vs. Frequency

4.7 Comparison With Previous Work

Table 4.1 is a comparison to the previous work in this field. Sowlati's device [8] offers good performance, but is designed in a technology that is difficult to integrate with baseband circuitry. Su's [9] and Doyama's [10] devices are both built in CMOS but are designed for the lower 830 MHz AMPS frequency band.

Tsai's device [5] offers good performance but has been improved in two particular areas. Firstly, Tsai uses off-chip input matching. In an actual implementation, an input matching network must be integrated along with the amplifier on-chip. Secondly, the architecture of the amplifier is such that the circuit self-oscillates. This results in a circuit that requires a large input drive power to stabilize. The resultant circuit will also occupy greater die space than would otherwise be necessary.

Table 4.1: Comparison to Previous Work

	Sowlati[8]	Su[9]	Tsai[5]	Doyama[10]	This Work
Year	1996	1997	1998	1999	-
Technology	0.8 μm GaAs MESFET	0.8 μm CMOS	0.35 μm CMOS	0.35 μm CMOS	0.35 μm CMOS
Frequency	1.9 GHz	824 – 829 MHz	1.9 GHz	835 MHz	1.9 GHz
Supply Voltage	2.4 V	2.5 V	1 V	3.3 V	2.5
Efficiency	57%	42%	40%	36%	38%
Power Gain	16 dB	25 dB	-	-	13 dB
Output Power	200 mW	1000 mW	200 mW	79 mW	185 mW
Die Area	5.7 mm ²	1.5 mm ²	0.8 mm ²	1.9 mm ²	1.4 mm ²
Architecture	single ended	single ended	differential	single ended	differential

4.8 Conclusions

In this chapter, a number of testing techniques were described. Considerations necessary for both low and high power testing were also examined. High power testing entailed the design of a specialized test fixture capable of handling the RF currents and thermal dissipation required for this design. Power testing revealed a 38 % power added efficiency with a 22.5 dBm power output.

4.9 Reference

- 1 T. Ruttan, "Designing amplifiers for wireless systems." *Microwaves & RF*, January, pp. 89-100, 1998.
- 2 K. Jones, "Low-voltage RF IC amps drive handset applications." *Microwaves & RF*, February, pp. 141-146, 1998.
- 3 W. Titus, R. Croughwell, C. Schiller, L. DeVito, "A SI BJT RF dual band receiver IC for DAB." *IEEE Radio Frequency Integrated Circuits Symposium*, vol. 3, pp. 297-300, 1998.
- 4 C. K. Lee, C. C. Ku, K. L. Su, C. H. Lin, K. C. Tao, "A 900 MHz ISM band transceiver RF IC chip set and RF module." *IEEE Radio Frequency Integrated Circuits Symposium*, vol. 1, pp. 245-248, 1998.
- 5 K. Tsai, P. R. Gray, "A 1.9 GHz, 1-W Class E power amplifier for wireless communications." *IEEE Journal of Solid State Circuits*, vol. 34, pp. 962-970, 1999.
- 6 M. Loy, M. Buschbom, M. Peterson, K. Nguyen, "Thermal design ensures RF power amp reliability." *Microwaves & RF*, July, pp. 55-65, 1998.
- 7 S. J. Maeng, S. S. Chun, J. L. Lee, C. S. Lee, K. J. Youn, D. M. Park, "A GaAs power amplifier module for 3.3 V CDMA/AMPS Dual Mode Cellular Phones." *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, pp. 2839-2843, 1995.
- 8 T. Sowlati, Y. Greshishchev, and C.A.T. Salama "1.8 GHz Class E Power Amplifier for Wireless Communications," *Electronic Letters*, vol. 32, pp. 1846-1848, 1996.
- 9 D. Su and W. McFarland, "A 2.5 V, 1 W Monolithic CMOS RF Power Amplifier." *IEEE Custom Integrated Circuits Conference. Proceedings*, 1997.
- 10 J. Doyama, "CMOS Class E Power Amplifier." M.A.Sc. Thesis, University of Toronto, 1999.

CHAPTER 5

Conclusions

This thesis presented the design of an RF amplifier for use in PCS band CDMA applications. The goal of this thesis was to design an RF amplifier capable of delivering 200 mW of power, with as high efficiency as possible. It was also desired to implement the amplifier in a conventional 0.35 μm CMOS process without the benefit of thinned dies or any other additional exotic processing steps. The fabricated chip occupied 1.4 mm^2 . The amplifier design utilized a differential topology and was implemented with an on-chip matching network.

The amplifier was initially tested at DC levels to validate functionality. Later, low power testing was conducted at the operating frequency to determine stability. A customized test fixture was developed to expand the high-power capability of the design. Finally, a load pull analysis was performed during testing to extract peak performance.

The amplifier performance was noted to be stable and free from any spurious signals. The input matching network functioned well and achieved a better than 12 dB match across the transmit band. The final output power was found to be 22.5 dBm at a 38 percent efficiency.

Future work should include modeling to calculate the parasitic losses associated with the chip architecture. A full-wave electromagnetic analysis should therefore be conducted to make these calculations.

It should be apparent that high power amplifiers require special implementation and testing considerations. The use of the chip-on-board method was therefore

implemented to address some of these concerns. Given the relative success of the chip-on-board technique, it is therefore highly recommended that this technique be the prime focus of future power amplifier designs. It is also recommended that future designs allow the die countersunk into the circuit board. This would permit the die to be directly mounted onto a large heat spreader. This approach would also allow for even shorter bondwires to be used.

It was the intention of this work to implement an amplifier using only high volume production techniques. There have been a vast number of works that have made use of exotic processing steps to produce highly beneficial results. Of the large number of new technologies discussed in this thesis, in general, those that would permit faster device transition times and a reduced drain capacitance are key.

Passive component performance will suffer with the increasingly less resistive substrates needed for future technologies. The optimal choice would therefore would be the use of silicon-on-insulator, SOI, technologies. In this case, the bulk field substrate is an insulating oxide. The use of SOI technologies will therefore minimize the capacitive losses associated with spiral inductors.